CSEDU 2014

6th INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED EDUCATION

PROCEEDINGS

VOLUME 2

1 - 3 APRIL, 2014

BARCELONA, SPAIN

WWW.CSEDU.ORG

SPONSORED BY: INSTICC

IN COOPERATION WITH: acm In-Cooperation, SIGCSE, European Council for Business Education, ATIEF

TECHNICALLY CO-SPONSORED BY: IEEE Sociedad Educación Capítulo Español, spee

MEDIA PARTNERS: Open Education Europa

CSEDU 2014

Proceedings of the
6th International Conference on
Computer Supported Education

Volume 2

Barcelona, Spain

1 - 3 April, 2014

Sponsored by
INSTICC – Institute for Systems and Technologies of Information, Control and Communication

In Cooperation with
ACM SIGITE – Association for Computing Machinery - Special Interest Group for Information Technology Education
ECBE – European Council for Business Education
ATIEF

Technically Co-sponsored by
IEEE Education Society
SPEE – Portuguese Society for Engineering Education

Media Partner
Open Education Europa
Internet Multimedia Search and Mining

Copyright © 2014 SCITEPRESS – Science and Technology Publications
All rights reserved

Edited by Susan Zvacek, Maria Teresa Restivo, James Uhomoibhi and Markus Helfert

Printed in Portugal

ISBN: 978-989-758-021-5

Depósito Legal: 372332/14

http://www.csedu.org

csedu.secretariat@insticc.org

BRIEF CONTENTS

INVITED SPEAKERS IV

ORGANIZING AND STEERING COMMITTEES V

PROGRAM COMMITTEE VI

AUXILIARY REVIEWERS XI

SELECTED PAPERS BOOK XI

FOREWORD XIII

CONTENTS XV

INVITED SPEAKERS

José Carlos Lourenço Quadrado

ISEL - Lisbon Superior Engineering Institute, Portugal and IFEES - International Federation of Engineering Education Societies

U.S.A.

Steve Wheeler

Plymouth Institute of Education, Plymouth University

U.K.

Larissa Fradkin

London South Bank University, Brunel University and Sound Mathematics Ltd.

U.K.

Erik de Graaff

Aalborg University

Denmark

ORGANIZING AND STEERING COMMITTEES

CONFERENCE CHAIR

Markus Helfert, Dublin City University, Ireland

PROGRAM CO-CHAIRS

Susan Zvacek, Fort Hays State University, U.S.A.

Maria Teresa Restivo, FEUP, Portugal

James Uhomoibhi, University of Ulster, U.K.

PROCEEDINGS PRODUCTION

Marina Carvalho, INSTICC, Portugal

Helder Coelhas, INSTICC, Portugal

Ana Guerreiro, INSTICC, Portugal

Filipe Mariano, INSTICC, Portugal

Andreia Moita, INSTICC, Portugal

Raquel Pedrosa, INSTICC, Portugal

Vitor Pedrosa, INSTICC, Portugal

Cláudia Pinto, INSTICC, Portugal

Cátia Pires, INSTICC, Portugal

Ana Ramalho, INSTICC, Portugal

Susana Ribeiro, INSTICC, Portugal

Rui Rodrigues, INSTICC, Portugal

Sara Santiago, INSTICC, Portugal

André Santos, INSTICC, Portugal

Fábio Santos, INSTICC, Portugal

José Varela, INSTICC, Portugal

CD-ROM PRODUCTION

Pedro Varela, INSTICC, Portugal

GRAPHICS PRODUCTION AND WEBDESIGNER

André Lista, INSTICC, Portugal

Mara Silva, INSTICC, Portugal

SECRETARIAT

Andreia Pereira, INSTICC, Portugal

WEBMASTER

Susana Ribeiro, INSTICC, Portugal

PROGRAM COMMITTEE

Mehdi Adda, Université du Québec à Rimouski, Canada

Shirley Agostinho, University of Wollongong, Australia

Carlos Alario-Hoyos, Universdad Carlos III, Spain

Peter Albion, University of Southern Queensland, Australia

Nelma Albuquerque, People-Inspired Technologies, The Netherlands

Efthimios Alepis, University of Piraeus, Greece

Vassil Alexandrov, Barcelona Supercomputing Center, Spain

Colin Allison, University of St. Andrews, U.K.

Marilisa Amoia, Saarland University, Germany

George Anastassakis, University of Piraeus, Greece

Terry Anderson, Athabasca University, Canada

Galia Angelova, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Bulgaria

Leonard Annetta, George Mason University, U.S.A.

Juan Ignacio Asensio, University of Valladolid, Spain

Michael E. Auer, Carinthia Tech Institute, Austria

Anders Avdic, Örebro University, Sweden

Breno Fabrício Terra Azevedo, Instituto Federal de Educação, Ciência e Tecnologia Fluminense, Brazil

Paul M. A. Baker, Georgia Institute of Technology, U.S.A.

Adriano Baratè, Università degli Studi di Milano, Italy

Jorge Barbosa, UNISINOS, Brazil

Patrícia Brandalise Scherer Bassani, Universidade Feevale, Brazil

Clemens Bechter, TBS, Thailand

Patricia Alejandra Behar, Universidade Federal do Rio Grande do Sul, Brazil

Kay Berkling, Cooperative State University Baden Württemberg, Karlsruhe, Germany

Marlies Bitter-Rijpkema, Open Universiteit, The Netherlands

Emmanuel G. Blanchard, University of Montreal, France

Lars Bollen, Universiteit Twente, The Netherlands

Andreas Bollin, Alpen-Adria Universität Klagenfurt, Austria

Katrin Borcea-Pfitzmann, Technische Universität Dresden, Germany

Federico Botella, Miguel Hernandez University of Elche, Spain

Jacqueline Bourdeau, Télé-université, UQAM, Canada

Ghizlane El Boussaidi, École de Technologie Supérieure, Canada

Patrice Bouvier, Université de Lyon, France

Claudio da Rocha Brito, COPEC, Brazil

Dumitru Burdescu, University of Craiova, Romania

Santi Caballé, Open University of Catalonia, Spain

Chris Campbell, University of Queensland, Australia

Alberto Cardoso, University of Coimbra, Portugal

Carlos Vaz de Carvalho, ISEP, Portugal

Chia-Hu Chang, National Taiwan University, Taiwan

Amitava Chatterjee, Jadavpur University, India

Weiqin Chen, University of Bergen, Norway

Line Clemmensen, Technical University of Denmark, Denmark

Jo Coldwell-Neilson, Deakin University, Australia

Marc Conrad, University of Bedfordshire, U.K.

Gennaro Costagliola, Università di Salerno, Italy

Maria Cristina Costa-Lobo, Universidade Portucalense, Portugal

PROGRAM COMMITTEE (CONT.)

Manuel Perez Cota, Universidade de Vigo, Spain

Caroline M. Crawford, University of Houston-Clear Lake, U.S.A.

Beno Csapó, University of Szeged, Hungary

John Philip Cuthell, Virtual Learning, U.K.

Mats Daniels, Uppsala University, Sweden

Thanasis Daradoumis, University of the Aegean / Open University of Catalonia, Greece

Sergiu Dascalu, University of Nevada, Reno, U.S.A.

Stavros N. Demetriadis, Aristotle University of Thessaloniki, Greece

John Dempsey, University of South Alabama, U.S.A.

Giuliana Dettori, Istituto per le Tecnologie Didattiche (ITD-CNR), Italy

Alicia Diaz, La Plata Nacional University, Argentina

Yannis Dimitriadis, School of Telecommunications Engineering, University of Valladolid, Spain

Danail Dochev, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Bulgaria

Toby Dragon, Ithaca College, U.S.A.

Amalia Duch-Brown, UPC - Politechnical University of Catalonia, Spain

Ishbel Duncan, University of St. Andrews, U.K.

Erik Duval, Katholieke Universiteit Leuven, Belgium

Olga Dziabenko, Universidad de Deusto, Spain

Martin Ebner, Technische Universität Graz, Austria

Paula Escudeiro, Instituto Superior de Engenharia do Porto, Portugal

Barbara Di Eugenio, University of Illinois at Chicago, U.S.A.

Si Fan, University of Tasmania, Australia

Mingyu Feng, SRI International, U.S.A.

Richard E. Ferdig, Kent State University, U.S.A.

Davide Fossati, Carnegie Mellon University in Qatar (CMU- Qatar), U.S.A.

Rita Francese, Università degli Studi di Salerno, Italy

José Ricardo Franco, Universidade Federal de Minas Gerais, Brazil

Jun Fujima, Fraunhofer IDMT, Germany

Serge Garlatti, TELECOM Bretagne, France

Isabela Gasparini, UFRGS and UDESC, Brazil

Xun Ge, University of Oklahoma, U.S.A.

Sébastien George, Université du Maine, France

Henrique Gil, Escola Superior de Educação do Instituto Politécnico de Castelo Branco, Portugal

Denis Gillet, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Ana González-Marcos, Universidad de la Rioja, Spain

Anandha Gopalan, Imperial College London, U.K.

Grami Grami, King Abdulaziz University, Saudi Arabia

Angela Guercio, Kent State University, U.S.A.

Christian Guetl, Graz University of Technology, Austria

Matej Guid, Faculty of Computer and Information Science, University of Ljubljana, Slovenia

David Guralnick, Kaleidoscope Learning / Columbia University, U.S.A.

Laurence Habib, Oslo and Akerhus University College of Applied Sciences, Norway

Leontios Hadjileontiadis, Aristotle University of Thessaloniki, Greece

Ibrahim A. Hameed, Aalborg University, Denmark

Yasunari Harada, Waseda University, Japan

Andreas Harrer, Technical University Clausthal, Germany

Roger Hartley, University of Leeds, U.K.

PROGRAM COMMITTEE (CONT.)

Markus Helfert, Dublin City University, Ireland

Richard Helps, Brigham Young University, U.S.A.

George Hloupis, Technological Educational Institute of Athens, Greece

Mark van 't Hooft, Kent State University, U.S.A.

Janet Hughes, University of Dundee, U.K.

Kin-chuen Hui, Chinese University of Hong Kong, Hong Kong

Nataliya V. Ivankova, University of Alabama at Birmingham, U.S.A.

Ivan Ivanov, SUNY Empire State College, U.S.A.

Romina Jamieson-Proctor, University of Southern Queensland, Australia

Marc Jansen, University of Applied Sciences Ruhr West, Germany

Stéphanie Jean-Daubias, Université Claude Bernard - Lyon 1 / LIRIS, France

Lars Johnsen, University of Southern Denmark, Denmark

Katerina Kabassi, Tei of the Ionian Islands, Greece

Michail Kalogiannakis, University of Crete, Greece

Achilles Kameas, Hellenic Open University, Greece

Charalampos Karagiannidis, University of Thessaly, Greece

Ilias Karasavvidis, University of Thessaly, Greece

Göran Karlsson, KTH Stockholm, Sweden

David Kaufman, Simon Fraser University, Canada

Jalal Kawash, University of Calgary, Canada

ChanMin Kim, The University of Georgia, U.S.A.

John Kinnebrew, Vanderbilt University, U.S.A.

Dimitris Kiritsis, EPFL, Switzerland

Gwendolyn Kolfschoten, Delft University of Technology, The Netherlands

Rob Koper, Open University of the Netherlands, The Netherlands

Maria Kordaki, University of the Aegean, Greece

Miroslav Kulich, Czech Technical University in Prague, Czech Republic

Chronis Kynigos, University of Athens and CTI, Greece

Jean-Marc Labat, Université Pierre et Marie Curie, France

Timo Lainema, Turku School of Economics, Finland

H. Chad Lane, University of Southern California, U.S.A.

Reneta Lansiquot, New York City College of Technology, U.S.A.

Rynson Lau, City University of Hong Kong, Hong Kong

José Paulo Leal, Faculty of Sciences of the University of Porto, Portugal

Dominique Leclet, UPJV, France

Newton Lee, Newton Lee Laboratories LLC, U.S.A.

Andrew Lian, Suranaree University of Technology, Thailand

Yu-Tzu Lin, National Taiwan Normal University, Taiwan

Andreas Lingnau, Catholic University Eichstaett-Ingolstadt, Germany

Luca Andrea Ludovico, Università degli Studi di Milano, Italy

Heide Lukosch, Delft University of Technology, The Netherlands

Stephan Lukosch, Delft University of Technology, The Netherlands

Philip Machanick, Rhodes University, South Africa

Krystina Madej, Georgia Tech, U.S.A.

Massimo Marchiori, University of Padua, Italy

Jacek Marciniak, Adam Mickiewicz University, Poland

Ivana Marenzi, Leibniz University Hannover, Germany

PROGRAM COMMITTEE (CONT.)

José Couto Marques, FEUP, Portugal

Lindsay Marshall, Newcastle University, U.K.

Alke Martens, Universität of Rostock, Germany

Manolis Mavrikis, London Knowledge Lab, U.K.

Bruce Maxim, University of Michigan-Dearborn, U.S.A.

Roger McDermott, Robert Gordon University, U.K.

Rory McGreal, Athabasca University, Canada

Hamid Mcheick, University of Quebec at Chicoutimi, Canada

Bruce McLaren, Carnegie Mellon University, United States, Germany

António José Mendes, Universidade de Coimbra, Portugal

Robert Meolic, University of Maribor, Slovenia

José Carlos Metrôlho, Instituto Politécnico de Castelo Branco, Portugal

Bakhtiar Mikhak, Harvard University, U.S.A.

Alexander Mikroyannidis, The Open University, U.K.

Felix Mödritscher, Vienna University of Economics and Business, Austria

Gyöngyvér Molnár, University of Szeged, Hungary

Roberto Moriyón, Universidad Autonoma de Madrid, Spain

Jack Mostow, Carnegie Mellon University, U.S.A.

Antao Moura, Federal Universisty of Campina Grande (UFCG), Brazil

Muhanna Muhanna, Princess Sumaya University for Technology, Jordan

Jogesh K. Muppala, Hong Kong University of Science and Technology, Hong Kong

Hiroyuki Nagataki, Okayama University, Japan

Ryohei Nakatsu, National University of Singapore, Singapore

Minoru Nakayama, Tokyo Institute of Technology, Japan

Sotiris Nikolopoulos, Technological Educational Institute of Larissa, Greece

Ebba Ossiannilsson, Lund University, Sweden

José Palma, Escola Superior de Tecnologia de Setúbal, Portugal

Viktoria Pammer, Know-Center, Austria

Pantelis Papadopoulos, Carnegie Mellon University in Qatar, Qatar

Kyparisia A. Papanikolaou, School of Pedagogical & Technological Education, Greece

Iraklis Paraskakis, South East European Research Centre, Greece

Seung Won Park, Texas A&M University - Corpus Christi, U.S.A.

Mar Pérez-Sanagustín, Universidad Carlos III de Madrid, Spain

Antonio Ramón Bartolomé Pina, Universitat de Barcelona, Spain

Niels Pinkwart, Humboldt University, Germany

Leonard A. Plugge, Stichting SURF, The Netherlands

Elvira Popescu, University of Craiova, Romania

Philippos Pouyioutas, University of Nicosia, Cyprus

Franz Puehretmair, Competence Network Information Technology to Support the Integration of People with Disabilities (KI-I), Austria

Khalid Rabayah, Arab American University, Palestinian Territory, Occupied

Maria da Graça Rasteiro, University of Coimbra, Portugal

Janet C. Read, University of Central Lancashire, U.K.

Petrea Redmond, University of Southern Queensland, Australia

Manuel José Cabral Dos Santos Reis, University of Trás-os-montes e Alto Douro, Portugal

Manuela Repetto, Consiglio Nazionale delle Ricerche, Italy

Maria Teresa Restivo, FEUP, Portugal

PROGRAM COMMITTEE (CONT.)

Fernando Reinaldo Ribeiro, Instituto Politécnico de Castelo Branco, Portugal

Martin Rich, Cass Business School, U.K.

Manuela Raposo Rivas, Universidad de Vigo, Spain

Razvan Rughinis, University "Politehnica" of Bucharest, Romania

Barbara Sabitzer, Alpen-Adria-Universität Klagenfurt, Austria

Libuse Samkova, University of South Bohemia, Czech Republic

Demetrios Sampson, University of Piraeus and Center For Research and Development - Hellas, Greece

Ian Sanders, University of South Africa, South Africa

Albert Sangra, Open University of Catalonia, Spain

Giuseppe Scanniello, Università Degli Studi della Basilicata, Italy

Sabine Schlag, University of Wuppertal, Germany

Ulrik Schroeder, RWTH Aachen University, Germany

Edgar Seemann, Furtwangen University, Germany

Pei Hwa Siew, Universiti Tunku Abdul Rahman, Malaysia

Martin Sillaots, Tallinn University, Estonia

Juarez Bento da Silva, Universidade Federal de Santa Catarina, Brazil

Peter Sloep, Open University of the Netherlands, The Netherlands

Zacharoula Smyrnaiou, University of Athens, Greece

Filomena Soares, University of Minho, Portugal

Michael Sonntag, Johannes Kepler University, Austria

J. Michael Spector, University of North Texas, U.S.A.

Karl Steffens, University of Cologne, Germany

Craig Stewart, Indiana University, U.S.A.

Hongyi Sun, City University of Hong Kong, Hong Kong

Katsuaki Suzuki, Kumamoto University, Japan

Ekaterina Talalakina, NRU HSE, Russian Federation

Bénédicte Talon, ULCO, Université Lille Nord de France, France

Brendan Tangney, Trinity College Dublin, Ireland

Steven Tanimoto, University of Washington, U.S.A.

Pei-Lee Teh, Monash University, Malaysia

Uwe Terton, University of the Sunshine Coast, Australia

Neena Thota, University of Saint Joseph, Macau

Fatt Hee Tie, University of Malaya, Malaysia

Thanassis Tiropanis, University of Southampton, U.K.

Stefan Trausan-Matu, University Politehnica of Bucharest, Romania

Joseph Trimmer, Ball State University, U.S.A.

George Tsihrintzis, University of Piraeus, Greece

James Uhomoibhi, University of Ulster, U.K.

Carsten Ullrich, Shanghai Jiao Tong University, China

Luis de la Fuente Valentín, Universidad Internacional de la Rioja, Spain

Andreas Veglis, Aristotle University of Thessaloniki, Greece

Ioanna Vekiri, Independent Researcher, Greece

Maria Virvou, University of Piraeus, Greece

Harald Vranken, Open Universiteit, The Netherlands

Pieter de Vries, Delft University of Technology, The Netherlands

Minjuan Wang, San Diego State University, U.S.A.

Martin Wessner, Fraunhofer IESE, Germany

Su White, Web and Internet Science, U.K.

PROGRAM COMMITTEE (CONT.)

Leandro Krug Wives, Universidade Federal do Rio Grande do Sul, Brazil

Jie Chi Yang, National Central University, Taiwan

Takami Yasuda, Nagoya University, Japan

Katarina Zakova, Slovak University of Technology in Bratislava, Slovak Republic

Diego Zapata-Rivera, Educational Testing Service, U.S.A.

Fani Zlatarova, Elizabethtown College, U.S.A.

Amal Zouaq, Royal Military College of Canada, Canada

Javier García Zubía, Universidad de Deusto, Spain

Susan Zvacek, Fort Hays State University, U.S.A.

AUXILIARY REVIEWERS

Alvine Boaye Belle, Ecole de Technologie Superieure, Canada

José Cordeiro, Polytechnic Institute of Setúbal / INSTICC, Portugal

Nuno Pina Gonçalves, EST-Setúbal / IPS, Portugal

Drew Hicks, NC State University, U.S.A.

Timotej Lazar, Univerza v Ljubljani, Fakulteta za racunalništvo in informatiko, Slovenia

María del Mar Sánchez-Vera, University of Murcia, Spain

Jian Shi, University of Southampton, U.K.

Yasuhisa Tamura, Sophia University, Japan

Robert Thompson, University of Washington, U.S.A.

SELECTED PAPERS BOOK

A number of selected papers presented at CSEDU 2014 will be published by Springer-Verlag in a CCIS Series book. This selection will be done by the Conference Chair and Program Co-chairs, among the papers actually presented at the conference, based on a rigorous review by the CSEDU 2014 Program Committee members.

FOREWORD

This book contains the proceedings of the 6th International Conference on Computer Supported Education (CSEDU 2014) which was organized and sponsored by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC). This conference was held in Cooperation with the Association for Computing Machinery - Special Interest Group for Information Technology Education, European Council for Business Education (ECBE) and Association des Technologies de l'Information pour l'Education et la Formation (ATIEF). The conference was also technically co-sponsored by IEEE Education Society - Capitulo Español and SPEE (Sociedade Portuguesa para a Educação em Engenharia).

CSEDU has become an annual meeting place for presenting and discussing learning paradigms, best practices and case studies that concern innovative computer-supported learning strategies, institutional policies on technology-enhanced learning including learning from distance, supported by technology. The Web is currently a preferred medium for distance learning and the learning practice in this context is usually referred to as e-learning or technology-enhanced learning. CSEDU 2014 provided an overview of the state of the art in technology-enhanced learning and outlined upcoming trends and promoting discussions about the educational potential of new learning technologies in the academic and corporate world.

This conference brought together researchers and practitioners interested in methodologies and applications related to the education field. It had five main topic areas, covering different aspects of Computer Supported Education, including "Information Technologies Supporting Learning", "Learning/Teaching Methodologies and Assessment", "Social Context and Learning Environments", "Domain Applications and Case Studies" and "Ubiquitous Learning". We believe these proceedings demonstrate new and innovative solutions, and highlight technical problems in each field that are challenging and worthwhile.

CSEDU 2014 received 242 paper submissions from 57 countries in all continents. A double-blind review process was enforced, with the help of the 277 experts who were members of the conference program committee, all of them internationally recognized in one of the main conference topic areas. Only 32 papers were selected to be published and presented as full papers, (i.e., completed work, 12 pages in proceedings, 30-minute oral presentations) for a 13% acceptance rate. Another 69 papers, describing work-in-progress, were selected as short papers for 20' oral presentation and 71 papers were presented as posters. All accepted papers represent a high level of quality, which we intend to maintain and reinforce in the next edition of this conference.

The high quality of the CSEDU 2014 programme was enhanced by four keynote lectures, delivered by experts in their fields, including (alphabetically): Erik de Graaff (Aalborg University, Denmark), José Carlos Lourenço Quadrado (ISEL - Lisbon Superior Engineering Institute, Portugal and IFEES - International Federation of Engineering Education Societies, United States), Larissa Fradkin (London South Bank University, Brunel University

and Sound Mathematics Ltd., United Kingdom) and Steve Wheeler (Plymouth Institute of Education, Plymouth University, United Kingdom).

For the sixth edition of the conference we extended and ensured appropriate indexing of the proceedings of CSEDU including Thomson Reuters Conference Proceedings Citation Index, INSPEC, DBLP, EI and Scopus. Besides the proceedings edited by SCITEPRESS, a short list of papers presented at the conference will be selected so that revised and extended versions of these papers will be published by Springer-Verlag in a CCIS Series book. Furthermore, all presented papers will soon be available at the SCITEPRESS digital library.

The best contributions to the conference and the best student submissions were distinguished with awards based on the best combined marks of paper reviewing, as assessed by the Program Committee, and the quality of the presentation, as assessed by session chairs at the conference venue.

The conference was complemented with a special session, focusing on Well-being Literacy; named Special Session on Well-being Literacy through Multimedia Education for Vulnerable Populations - WeLL 2014.

Building an interesting and successful program for the conference required the dedicated effort of many people. Firstly, we must thank the authors, whose research and development efforts are recorded here. Secondly, we thank the members of the program committee and additional reviewers for their diligence and expert reviewing. We also wish to include here a word of appreciation for the excellent organization provided by the conference secretariat, from INSTICC, which has smoothly and efficiently prepared the most appropriate environment for a productive meeting and scientific networking. Last but not least, we thank the invited speakers for their invaluable contribution and for taking the time to synthesize and deliver their talks.

Markus Helfert
Dublin City University, Ireland

Susan Zvacek
Fort Hays State University, U.S.A.

Maria Teresa Restivo
FEUP, Portugal

James Uhomoibhi
University of Ulster, U.K.

CONTENTS

INVITED SPEAKERS

KEYNOTE SPEAKERS

Computer Supported Education - The Human Factor IS-5
José Carlos Lourenço Quadrado

Digital Age Learning - The Changing Face of Online Education IS-7
Steve Wheeler

Mathematics Teaching - Is the Future Syncretic? IS-9
Larissa Fradkin

Team Learning in Engineering Education IS-11
Erik de Graaff

LEARNING/TEACHING METHODOLOGIES AND ASSESSMENT

FULL PAPERS

Learning on Electrical Circuits While Playing 'E&E Electrical Endeavours' - Design Research on a Serious Game Optimizing for Conceptual Understanding 5
Ruurd Taconis, Mariette Dubois, Lesley de Putter and Henry van Bergen

Automated Generation of Geometry Questions for High School Mathematics 14
Rahul Singhal, Martin Henz and Kevin McGee

Finding Regularities in Courses Evaluation with K-means Clustering 26
R. Campagni, D. Merlini and M. C. Verri

Cross-analysis of Transversal Competences in Project Management 34
Ana González-Marcos, Fernando Alba-Elías, Joaquín Ordieres-Meré and Fermín Navaridas-Nalda

Getting it Fast - An Information Systems Oriented Semantic Web Curriculum 42
Daniela Giordano and Francesco Maiorana

Teaching Web Programming - An Approach Rooted in Database Principles 49
Francesco Maiorana

The Elevate Framework for Assessment and Certification Design for Vocational Training in European ICT SMEs 57
Iraklis Paraskakis and Thanos Hatziapostolou

Integrated STEM in Elementary Grades Using Distributed Agent-based Computation 67
Pratim Sengupta, Gokul Krishnan and Mason Wright

SHORT PAPERS

Layered Evaluation Based on Timed Automata for Educational Scenarios - Case Study on Blended Learning 81
Manel Bensassi and Mona Laroussi

Hints for Organizing a Successful Doctoral Symposium - An Experience Report 89
Claudia Raibulet and Eila Ovaska

Adopting Commercially Inspired Practices Within an Academic Teaching Course - A Case Study of a Computer Games Engineering Degree 95
G. Ushaw, W. Blewitt and G. Morgan

Continuous Assessment in Civil Engineering Education - Yes, but with Some Conditions 103
F. J. Martín-Carrasco, A. Granados, D. Santillan and L. Mediero

Game-based Learning in IT Service Transition - The Case of a Mobile Sales Service by a Small Team in Brazil 110
Thiago Paiva Brito, Josias Paes and J. Antão B. Moura

Bringing Real-life Practice in Software Project Management Training Through a Simulation-based Serious Game 117
Alejandro Calderón and Mercedes Ruiz

Blended Learning as a Strategic Method Against the Illegal Use of Internet 125
Julija Lapuh Bele, Andreja Sladoje Jemec, David Rozman and Maja Dimc

Data Model and Smartphone App in an Observational Research Social Network 131
Cristina Mesquita and Rui Pedro Lopes

Automatic Objective Assessments of Japanese Reading Difficulty with the Operation Records on Japanese Text Presentation System - Design and Implementation 139
Kyota Aoki, Shinjiro Murayama and Kohji Harada

Review of Learner Modeling Using Educational Games 147
Mohamed Ali Khenissi, Fathi Essalmi, Mohamed Jemni and . Kinshuk

Detecting Math-and-ICT Competence 153
Maria Shabanova and Borislav Lazarov

The Life Cycle of a Cutting-edge Technology Course - A Coaching Experience on Android 159
Maria J. Blesa, Amalia Duch, Joaquim Gabarró and María J. Serna

Developed and Taught Course Modules to Enhance Cloud Computing Education 167
Huiming Yu, Nakia Powell, Dexter Stembridge and Jonathan Rann

Better Feedback for Educational Online Judges 176
Anaga Mani, Divya Venkataramani, Jordi Petit and Salvador Roura

A Technology Enhanced Assessment System for Skill and Knowledge Learning 184
Enosha Hettiarachchi, Enric Mor, Maria Antonia Huertas and M. Elena Rodríguez

Developing Web-based Tutorial Modules to Support Training for Group-based Mentoring 192
Sharon Locke, Georgia Bracey and Stephen Marlette

A Problem-based Learning Case Study for Teaching Voice over Internet Protocol - VoIP - Using Asterisk as a Tool for Teaching VoIP for Information Technology Classes 197
M. C. Dias, C. F. Gabi, E. P. Rodrigues, V. R. Souza and A. Perkusich

Partial Correctness and Continuous Integration in Computer Supported Education 205
Daniela Fonte, Ismael Vilas Boas, Nuno Oliveira, Daniela da Cruz, Alda Lopes Gançarski and Pedro Rangel Henriques

The HumanoidLab - Involving Students in a Research Centre Through an Educational Initiative 213
Guillem Alenyà, José Luis Rivero, Aleix Rull, Patrick Grosch and Sergi Hernández

Student Course Evaluation - Class Size, Class Level, Discipline and Gender Bias 221
Jacob Kogan

KASS: Korean Automatic Scoring System for Short-answer Questions 226
Eun-Seo Jang, Seung-Shik Kang, Eun-Hee Noh, Myung-Hwa Kim, Kyung-Hee Sung and Tae-Je Seong

A Blended Health Science Education Graduate Program - Integrating Technology Assignments into Coursework 231
Sharla King, Kim Peacock and Lili Liu

Using Self-service Business Intelligence for Learning Decision Making with Business Simulation Games 235
Waranya Poonnawat and Peter Lehmann

MOOC and Mechanized Grading 241
Christian Queinnec

Integrating Computers, Science, and Mathematics - A Course for Future Mathematics Teachers 246
Alfinio Flores

Learning Support System for Paleontological Environment Based on Body Experience and Sense of Immersion - Extinct Animals Move in Synchronization with Human Actions 252
Tomohiro Nakayama, Kaori Izuishi, Fusako Kusunoki, Ryuichi Yoshida, Takayuki Adachi, Takeki Ogitsu, Hiroshi Takemura, Hiroshi Mizoguchi and Shigenori Inagaki

Multiple Line System for Visualizing and Sharing Students' Commitment on the Tablet PC 258
Keita Muratsu, Hideo Funaoi, Etsuji Yamaguchi and Shigenori Inagaki

The Virtual Classroom - A Pilot Case in Inquiry Based Learning 264
Pavel Boytchev, Eliza Stefanova, Nikolina Nikolova and Krassen Stefanov

STEM Education - Students Touch Space Through Free Robotics Programming Competition 270
Jenny Liu, Wendy Feenstra, Alvar Saenz Otero and Kathleen Magrane

An Alternate Learning Approach for Destructive Testing in Civil Engineering - Benefits from Remote Laboratory Experimentation 275
Xuemei Liu, Nannan Zong and Manicka Dhanasekar

How do Student Evaluations of Courses and of Instructors Relate? 280
Tamara Sliusarenko, Line H. Clemmensen and Bjarne Kjær Ersbøll

Reflections on the Use of a Personal Response System (PRS) for Summative Assessment in an Undergraduate Taught Module 288
Alan Hilliard

Two New Concepts in Video Podcasts - Minimalist Slides and Modular Teaching Mini-videos 292
Emilio Letón and Elisa M. Molanes-López

Academic Frontier-based Approach Based on Constructivism 298
Nuo Liu, Shengli Yang, Bing Feng and Liangyin Chen

Using Technology to Accelerate the Construction of Concept Inventories - Latent Semantic Analysis and the Biology Concept Inventory 301
Kathy Garvin-Doxas, Michael Klymkowsky, Isidoros Doxas and Walter Kintsch

The Image of Images as an Aid to Improve Learning - An Eye-tracking Experiment Studying the Effect of Contrasts in Computer-based Learning Material 309
Mona Holmqvist Olander, Eva Wennås Brante and Marcus Nyström

Mediator-learner Dyad - Cooperative Relationships 317
Gilda Helena Bernardino de Campos and Gianna Oliveira Bogossian Roque

Using Activity Theory in Developing Instructial Tools for Project Management Studies 323
Maritta Pirhonen

Transforming Experience of Computer Science Software Development Through Developing a Usable Multiplayer Online Game in One Semester 329
Ilmi Yoon and Eun-young Elaine Kang

r(e)flect - The Reflective Teaching Material about Energy, Behaviour and Product Development 336
Magdalena Boork, Susanne Engström, Rebekah Olsen and Therese Balksjö

Supporting Creative Design Processes for the Support of Creative Mathematical Thinking - Capitalising on Cultivating Synergies between Math Education and Environmental Education 342
Chronis Kynigos and Maria Daskolia

DOMAIN APPLICATIONS AND CASE STUDIES

FULL PAPER

Fostering Information Literacy in German Psychology Students Using a Blended Learning Approach 353
Nikolas Leichner, Johannes Peter, Anne-Kathrin Mayer and Günter Krampen

SHORT PAPERS

A Study on the Permeation and Scope of ICT Intervention at the Indian Rural Primary School Level 363
Shrabastee Banerjee, Kalyan Sankar Mandal and Priyadarshini Dey

"Smoking Does Not Make You Happy" - Unlearning Smoking Habits Through Mobile Applications on Android OS 371
Răzvan Rughinis, Stefania Matei and Cosima Rughinis

3D Online Virtual Museum as e-Learning Tool - A Mixed Reality Experience 379
Dragoş Gheorghiu and Livia Ştefan

Envisioning an Advanced ICT-supported Build-up of Manufacturing Skills for the Factories of the Future 389
Joao Costa, Dimitris Kiritsis, Poul Kyvsgaard Hansen, Manuel Oliveira, Loukas Rentzos, Hadrien Szigeti and Marco Taisch

Development and Practical Application of a Relationship Diagram-creation Tool Centering on Automatic Creation Functionality 394
Norio Ishii, Yuri Suzuki and Saori Sakuma

AUTHOR INDEX 403

INVITED SPEAKERS

KEYNOTE SPEAKERS

Computer Supported Education
The Human Factor

José Carlos Lourenço Quadrado
ISEL - Lisbon Superior Engineering Institute, Portugal
IFEES - International Federation of Engineering Education Societies, U.S.A.

Abstract: Within the Computer Supported Education (CSE) there are three interesting trends observed.
The first one is the conscience that the technical specialisms are becoming increasingly important. New technologies and the emergence of Mobile, Cloud, Continuous Integration & Deployment allow educators with practical hands-on knowledge and experience to be decisive for the success of CSE projects.
On the other hand, it has changed the way we work together with each other. The necessity to have agile replies to the demands of the education stakeholders, to communicate and take responsibility, demands the critical thinking, the exchange of information and the flexibility being more and more the decisive success factors. These stakeholders are increasing creating a CSE team that consists of individuals more than ever with their own specialty and own drivers.
In addition, the time-to-market of the technology is getting shorter and is strongly dependent on the costs in the current economic climate.
All three trends converge to the manager. As a result, it is noticeable that you have to demonstrate the added value of your part in the CSE team, you must make results visible and under high pressure. How do you do that? Why is man such an important factor? What do you do with the man in your team? How do you put people in their strength and you create a work environment in which everyone comes into its own?

BRIEF BIOGRAPHY

José Carlos Quadrado is the full professor with tenure of electrical machines in the electrical engineering and automation department of the Instituto Superior de Engenharia de Lisboa (ISEL), Portugal.

Currently he holds the position of President of ISEL since 2006.

He has a BSc in Energy and Power Systems, a diploma degree in Electrical Engineering, Automation and Industrial Electronics from ISEL, a MSc and a Doctor degree in Electrical Engineering and Computers from Lisbon Technical University. He also holds the Habilitation degree (Aggregation) in Electrical Engineering from Beira Interior University.

Holds the position of President of the International Federation of Engineering Education Societies (IFEES) and the position of immediate-past president of the Ibero-American Engineering Education Association (ASIBEI), and he is also the immediate past vice-President of the European Society for Engineering Education (SEFI). Former member of the National Bologna Expert Group, he leads the Portuguese Observatory on European and Latin-American University management strategy best practices and the national association of engineering rectors and deans.

Being a member and senior member of several engineering societies and engineering education societies in several continents, he is also a visiting professor in several universities around the world and board member of technological societies.

He holds over 100 international publications (including journals and chapters of books), several patents and some international technical prizes and scholarships, and also held the position of editor and editor-in-chief in some journals. Up to now he has also developed several international engineering projects in the fields of renewable energy, fuel cells, electrical vehicles and intelligent control.

Digital Age Learning
The Changing Face of Online Education

Steve Wheeler
Plymouth Institute of Education, Plymouth University, U.K.

EXTENDED ABSTRACT

Online education was once quite simple. As a new form of distance education, it bridged the gap between those who could reach formal education and those, for whatever reason, who could not. Many institutions invested in web delivery as a primary means of engaging remote students, and a whole discourse rose up around the idea of 'online education.' It was a radical departure from the traditional correspondence courses, because it enabled instantaneous communication and feedback between students and their teachers, and allowed synchronous as well as asynchronous delivery of content. Video and live chat soon because desirable, and even essential facets of the online learning strategy. Content delivery was controlled by the experts through Learning Management Systems, and discussions were conducted via e-mail and bulletin boards. There was talk of the Virtual University, but it never really materialised. Instead, many universities adopted a blended (earlier referred to as a dual mode) delivery of their courses, which involved all of the above techniques, and a mix of other face-to-face and paper based activities. The online delivery model became a stable, reliable method for distance education for millions of students around the globe.

Then came the advent of Web 2.0 and social media and things changed, boundaries blurred and the pace of change accelerated. The rapid development of platforms and services that enabled users to both upload and download content was a departure from the previous, 'sticky' and content led approach of Web 1.0. Suddenly, in the age of the Social Web, user generated content was high profile and ubiquitous. Students began to share their own content through blogs, wikis, podcasts and videos. Platforms such as YouTube, WordPress, Blogger and Wikipedia became the first ports of call, and the user statistics grew exponentially. This was perplexing for many established online course providers. It was perplexing because they were no longer the sole arbiters of 'knowledge' and no longer held the exclusive right of mediators of learning. Many academics questioned the veracity of content that was created by 'amateurs', and many rejected that notion that knowledge could be generated, shared and repurposed freely. Issues of copyright, intellectual property and provenance became uppermost in the minds of teachers and lecturers. Many questions and much soul searching resulted:

In today's digital age, are Learning Management Systems still required, and is e-mail now increasingly anachronistic? Students now connect on social media tools such as Facebook and Twitter and converse through mobile text. They generate their own content on a regular basis and act as the nodes of their own production. Increasingly informal collaborating comes naturally through the new tools and technologies that constitute their personal learning environments. Add to this mix the meteoric success of Massive Online Open Courses and Flipped Classrooms, the popularity of online social games and prospect of new, emerging technologies such as augmented reality and wearable systems, and we continue to question the future of our current online educational provision. Is online learning now so common place and accessible that anyone can do it? Are we reaching the point in the development of technology that we are now in danger of repeating the mistakes we made when correspondence courses led to the ignominy of 'diploma mills' and ill reputed, fake degrees? What would we need to put into place to prevent this from happening in today's knowledge economy?

In this presentation I will address many of the above questions and offer my personal views on the future of online education in the Web 2.0 digital age. I will pose several questions including: What is

digital learning? How can teachers and students harness the power of tools and technologies for formal learning? What is the changing nature of knowledge? What will be the future of education, in the light of the radical changes and emerging technology? Ultimately, we will discuss issues that will impact upon all teachers and educators in the coming years: What will the next few years hold for online education? What will be the new reality for learning, knowledge and ultimately, human intelligence? Are learner expectations unfulfilled by the current provision of traditional educational institutions? How much will the roles of teachers be required to change? What new theories and practices will we need to develop to stay relevant in an increasingly technological world where the learner is taking control?

BRIEF BIOGRAPHY

Steve Wheeler is Associate Professor of Learning Technologies at Plymouth University, in South West England. Originally trained as a psychologist, he has spent his entire career working in media, technology and learning, predominantly in nurse education (NHS 1981-1995) and teacher education (1976-1981 and 1995-present). He is now in the School of Education, at the Faculty of Health, Education and Society.

Steve teaches on a number of undergraduate and post-graduate teacher education programmes in the UK and overseas. He researches into e-learning and distance education, with particular emphasis on the pedagogy underlying the use of social media and Web 2.0 tools, and he also has research interests in mobile learning and cyber-cultures. Steve is regularly invited to speak about his work and has given keynotes and invited lectures to audiences in 30 countries across 5 continents. He is currently involved in several research programmes related to e-learning, social media and handheld technologies.

Steve is the author of more than 150 scholarly articles, with over 2500 academic citations and is an active and prolific edublogger. His blog Learning with ‘e’s is a regular online commentary on the social and cultural impact of disruptive technologies, and the application of digital media in education and training. It currently attracts in excess of 150,000 views each month.

Steve is chair of the Plymouth e-Learning Conference, and between 2008-2011 was also co-editor of the journal Interactive Learning Environments. He serves on the editorial boards of a number of learning technology and education related open access academic journals including Research in Learning Technology (formerly ALT-J), the International Review of Research in Open and Distance Learning (IRRODL), the European Journal of Open, Distance and eLearning (EURODL) and Digital Culture and Education. He has served on the organising and executive committees of a number of international academic conferences, including ALT-C, ICL, EDEN, IFIP and AICT.

In 2008 Steve was awarded a Fellowship by the European Distance and E-learning Network (EDEN), and in 2011 he was elected to serve as a member of the Steering group of EDEN’s Network of Academics and Professionals (NAP). Between 2008-2013 he also served as chair of the influential worldwide research group IFIP Technical Committee Working Group 3.6 (distance education) and is author of several books including The Digital Classroom (Routledge: 2008) and Connected Minds, Emerging Cultures (Information Age: 2009). He lives in Plymouth, on the South West coast of England.

Mathematics Teaching
Is the Future Syncretic?

Larissa Fradkin
London South Bank University, Brunel University and Sound Mathematics Ltd., U.K.

EXTENDEND ABSTRACT

Teaching is one of the oldest professions on Earth and mathematics teaching must have had come first, since it appears that numbers had been invented before letters! Yet "mathematics wars" have been raging throughout the XX century and technology has only added fuel to fire: should mathematics be taught as poetry, requiring an inordinate amount of memorising and practising or should teachers concentrate on abstract concepts, with the tedium of calculations left to calculators and computers? Can ordinary learners grasp abstract concepts at all? On top of that, a modern University maths teacher teaching STEM students, particularly, future engineers has to cope with large classes, much larger than most European teachers had to deal with in the past. Can any of the teaching approaches be implemented in such environment in an effective manner? The advent of the XXI century saw mathematics teachers cajoled into employing the "evidence-based" technological solutions that had been shown to work when training University administrators, business managers or technicians. Many resisted, arguing that maths learning is a different process to learning a few words and procedures. Now it is all about Massive Online Open Courses and Flipped Classrooms. Can ordinary engineering students learn mathematics by watching MIT or Khan Academy videos? Can ordinary mathematics teachers facilitate the process by "flipping" in an effective way? I will present my thoughts on the subject, argue against false dichotomies and suggest syncretic solutions, including the ones that rely on cognitive technologies of the future.

I will start discussing several myths that became popular among educationalists:

Only what students discover for themselves is truly learned, while in reality students learn in a variety of ways. Recent research has shown that basing most learning on student discovery is time-consuming, does not insure that students end up learning the right concepts, and can delay or prevent progression to the next level. Successful programs use discovery for only a few very carefully selected topics, never all topics.

There are two separate and distinct ways to teach mathematics,… conceptual understanding through a problem solving approach and through drill and kill, while according to many experts success in mathematics needs to be grounded in well-learned algorithms as well as understanding of the concepts.

*Math concepts are best understood and mastered when presented "in context"; in that way, the underlying math concept will follow automatically, even though it has been repeatedly shown that w*hen story problems take centre stage, the math it leads to is often not practiced or applied widely enough for students to learn how to apply the concept to other problems.

I will move to recent MOOCs data that show that the pass rates of current MOOCs registrants are low and the overwhelming majority of those who obtain the certificate of achievement already have a degree. I will suggest my answers to the question posed by these data "What are MOOCs missing?" I will suggest that they open a door to advanced cognitive tutors that employ the oldest and best teaching method – Socratic dialogue, which allow students to master abstract concepts and practice under "expert supervision" at the time and place of their choosing.

Finally, I will address the suggestion that because of MOOCs in 50 years there will be only 10 institutions in the world delivering higher education. However, I would argue – in line with many other academics – that, particularly in the first year of their academic studies, students, disadvantaged STEM students need intensive guided teaching. This

can be delivered only in person and only by the most experienced teachers capable of not just "talking" – delivering content, but mainly of "listening" – and readjusting their delivery depending on the student immediate feedback. Of course, these experienced teachers and their students would benefit from access to quality MOOCs.

BRIEF BIOGRAPHY

Larissa Fradkin is Emerita Professor, London South Bank University, UK and Associated Professor, Brunel University, UK. Trained as a physicist at St Petersburg University, Russia, from 1974 till 1977 she studied for her PhD in Applied Mathematics at Victoria University of Wellington, New Zealand. From 1978 till 1984 she was employed as a Research Scientist at the NZ Department of Scientific and Industrial Research and from 1985 till 1992, as a Research Associate at the Department of Applied Mathematics and Theoretical Physics of Cambridge University, UK. From 1993 till 2009 Larissa was a member of academic staff at London South Bank University where she created and ran a Research Group on Mathematical Modelling of Ultrasonic NDE (Non-Destructive Evaluation of industrial components and materials) and teaching mathematics on various engineering courses. She is now a Managing Director of an independent research organisation Sound Mathematics Ltd., working on her ultrasonic projects, mainly in collaboration with CEA (the French Atomic Commission) and promoting sound mathematics education of engineers. Larissa is a Fellow of IMA (UK Institute for Mathematics and Its Applications), IOP (UK Institute of Physics) and IET (UK Institute of Engineering and Technology). Larisa has authored and co-authored over a 100 scientific and scholarly publications, lately on pedagogy behind her system of maths teaching and in 2013 Bookboon has published her open access e-textbook "College Algebra and Calculus: The Whys and Hows." Larissa has been a member of organising committees of several international engineering conferences and is currently a UK National contact person for SEFI's Maths Working Group (SEFI is the European Society for Engineering Education) and a member of the Committee on Mathematical Education for Engineers, iNEER (International Network for Engineering Education and Research).
Larissa lives in Cambridge, England.

Team Learning in Engineering Education

Erik de Graaff
Aalborg University, Denmark

Abstract: Most engineers work in teams during their professional life. Hence, learning to be an effective team member is an essential aspect of preparing for engineering practice. This presentation will analyse how Problem Based Learning (PBL) supports students in engineering in developing teamwork skills, like leadership, communication skills, and the ability to correctly assess your own contribution to the teamwork. Several tools that support this process will be discussed.

BRIEF BIOGRAPHY

Erik de Graaff is trained as psychologist and holds a PhD in social sciences. He has been working with Problem Based learning (PBL) in Maastricht from 1979 till 1990. In 1994 he was appointed as associate professor in the field of educational innovation at the Faculty of Technology Policy and Management of Delft University of Technology. Dr. de Graaff has been a visiting professor at the University of Newcastle, Australia in 1995 and a guest professor at Aalborg University in Denmark. The collaboration with Aalborg University led to an appointment as full professor at the department of Development and Planning in 2011. Dr. de Graaff is recognized as an international expert on PBL. He contributed to the promotion of knowledge and understanding of higher engineering education with numerous publications and through active participation in professional organizations like SEFI, IGIP, IFEES and ALE. He has published over 200 articles and papers and he has presented more than 70 keynotes and invited lectures on various topics related to PBL in higher education, like: Working with PBL, Management of change, Assessment and evaluation, Methods of applied research and Collaboration between university and industry. Since January 2008 he is Editor-in-Chief of the European Journal of Engineering Education.

LEARNING/TEACHING METHODOLOGIES AND ASSESSMENT

FULL PAPERS

Learning on Electrical Circuits While Playing 'E&E Electrical Endeavours'

Design Research on a Serious Game Optimizing for Conceptual Understanding

Ruurd Taconis[1], Mariette Dubois[2], Lesley de Putter[3,4] and Henry van Bergen[5]
[1]*Eindhoven School of Education, Eindhoven University of Technology, Den Dolech 2, Eindhoven, The Netherlands*
[2]*Sondervick College, Knegselseweg 30, Veldhoven, The Netherlands*
[3]*Heerbeeck College, Willem de Zwijgerweg 150, Best, The Netherlands*
[4]*Eindhoven University of Technology, Den Dolech 2, Eindhoven, The Netherlands*
[5]*Faculty of Electrical Engineering, Eindhoven University of Technology, Den Dolech 2, Eindhoven, The Netherlands*
r.taconis@tue.nl

Keywords: Design Research, Electrical Circuits, Conceptual Development, Serious Games.

Abstract: A serious game was developed in a two year design research project by educational researchers, game-designers and secondary school teachers in close collaboration. In a first round, students played the game in class using an open-inquiry strategy. Although the game had a strong impact on the students' conceptual development, it provoked the construction of misconceptions. The game was adapted and partially redesigned on the basis of the evaluation outcomes and an additional expert-review. Also the instructions to the game were redesigned and written down in a teachers' guide. In a second round, a pedagogical approach of alternating open-inquiry type gaming-episodes with guided reflection and internalisation episodes was used in class. Again a strong impact on students' conceptual understanding of electrical circuits was found. Significantly fewer misconceptions occurred. The results indicate that the close collaboration of school teachers, educational experts and game designers was fruitful for improving the serious game and its use in school practice. Moreover it became clear that serious games have the potential to contribute to students' conceptual understanding, in particular when a suitable mental model is coherently represented in the game´s layout and structure.

1 INTRODUCTION

Serious games are an inviting new option in education. The game industry is growing spectacularly. However, the characteristics that make games adequate *serious* games and the pedagogical do's and don'ts of using them in education are still largely uncovered (Michael and Chen, 2006). This paper reports on a design research project conducted using the game 'E&E electrical endeavours'. The project's aim is to help students master the subject of electrical circuits in grade 9 of Dutch secondary education, develop a pedagogical approach to using serious games in secondary education, and to motivate students for science and electrical engineering.

The game was previously developed at the faculty of Electrical Engineering at Eindhoven University of Technology mainly for raising the interest of potential new students. Hence the main focus in the initial version 1.0 of the E&E electrical endeavours game was on student involvement and motivation, excellent graphics and challenging content on electrical circuits. The game can be typified as a simulation based game for individual use. It is not a role-playing game where players are set against each other. In version 1.0 no match with formal science curricula was pursued. A pilot evaluation (Dekker et al., 2010) showed the game was attractive and entertaining for students in secondary school.

Secondary schools were seeking ways to modernize education by implementing ICT, to enrich science education by e.g. serious games, and to bring a lively perspective on scientific practice and careers into the school.

Combining these, the joint conclusion was that both university and secondary school goals could be successfully supported if an attractive serious game

could be constructed. The game should fit well into the formal curriculum, clearly contribute to student learning and should be brought into school as a sincere and attractive picture of scientific practice and careers.

To achieve this, a project group was formed comprising secondary school teachers, experts on science education and game-designers. Their task was to: "Construct and evaluate a serious game that is attractive and motivating on one hand and making an adequate contribution to formal learning in secondary education on the other". In line with this, a key research question addressed in this paper is: which characteristics of the game and of the way it is implemented in classroom make it adequate for learning about electrical circuits and dealing with misconceptions in particular? The approach taken to actually develop the game and answer the question was a design research project.

2 THEORY

2.1 Design Research

Design research is well known in both engineering and educational research (Gravemeijer & Cobb, 2006). In design research on education, education is being (re)designed and evaluated in various rounds. These comprise an alternating sequence of (re)design and evaluation leading to new or refined design criteria. The evaluation adds up to a picture that allows answering previously set research questions, as well as bringing forth new hypothesis and tentative answers. From a methodological point of view, it is a productive experimental design with both explorative and confirmative aspects. From a practitioner's point of view, it has the strong advantage of making a very close connection between research on one hand, and professional practice and professional development on the other.

2.2 Serious Games and Learning

Theory on the design and use of simulation-type serious games in education strongly emphasises the concepts of flow (Csikszentmihaly, 2009) and experiential learning. A key theory on experiential learning is developed by Kolb (1984), and adapted for the case of serious games by various authors (Ruben, 1999; Koops and Hoevenaar, 2012).

Using insights from general cognitive constructivist learning theories and the five stage model of skill acquisition by Dreyfus & Dreyfus (1980) (shown in Figure 1), Taconis (2011), developed a hypothetical model of experiential learning in serious games.



Figure 1: The Dreyfus & Dreyfus five stage model of the genesis of an expert (http://www.leanleadershipacademy.com/wp-content/uploads/2013/04/novice-expert1.jpg).

The 'Taconis model' explicitly pictures how the rules underlying and governing the game-engine, through producing the game's 'behaviour' and the regularities in the gaming environment, structure the learners experiences while gaming. It also pictures how skilled action by the learner, the acquisition of operational rules and ultimately a coherent reconstruction of the rules underlying the game-engine´s may result from the learners 'structured experiences'.



Figure 2: The hypothetical model of experiential learning in simulation-based serious games by Taconis (2011).

In the model 'skilled actions' (upper box in second column in Figure 2) are defined as those actions adequate within the gaming environment that do not imply or require conscious declarative knowledge.

These actions represent adequate behaviour in standard situations without being underpinned with knowledge. This corresponds to the *advanced beginner* position in the Dreyfus & Dreyfus five-stage-model.

'Operational rules' (middle box in second column in Figure 2) are – in line with the five- stage-model – considered a next step in development. The operational rules are typified by the involvement of conceptual understanding, though of an operational type, using concepts that are apparent within the experienced game environment. An example would be a learner that acts adequately within the game and is able to explain why and how he performs in terms of the game-world. This could be considered roughly equivalent to the stage of being *competent* within the five-stage-model though this concerns the game-environment only. Reaching this level requires the constructing (learning) of operational rules from experience, and it is presumed to require systematic and reflective thinking. Constructing these rules may be stimulated by asking the students questions, facilitating group discussions and adding tasks that require the construction and verbalization of such rules. Operational rules and skilled actions may transfer from the strict game-environment to other situations/environments, and this is more likely to occur the more akin these situations are (Vockell, 2013). For example, from a serious game that models projectile trajectories, skilled actions and operational rules may relatively easily transfer to projectile shooting experiments in classroom, but less easily to text-book questions on this subject or the context of satellite trajectories.

The (re)construction of 'theoretical insight' (lower box in second column in Figure 2), aligns with both '*competent* within the game-environment' towards '*competent* in real life'. This splitting-up of the 'competent stage' in the five-stage-model results from the fact that experiential learning from serious games involves two realities: the real world and the game world that simulates it. By its nature, theoretical insight will formally apply to real-life situations, since the rules a simulation based serious game is built upon apply to real world as well. Even though in many cases the rules underlying the game are in fact simplified versions of real life rules.

But this equivalence will not be automatically clear to the learner. Hence, theoretical insight only principally applies to both the game-world and the real world. In education, the learner should make considerable effort to understand and recognize this difference (transfer) and to achieve a 'competent in real-life' stage.

Constructing or reconstructing theoretical insight definitely requires systematic and reflective thinking as described above, but also making comparisons with contrasting and/or akin situations and theoretical points of view. The confrontation with situations other than the game-environment is needed; both to help constructing the theoretical ideas independently from the game-world and to facilitate transfer to other situations (Vockell, 2013).

The presumed necessity to implement in classthese reflective of theoretical components as well as exercises in other learning-environments than the game-environment, conflicts with the fundamental importance of flow for gaming and the use of serious games in education. Put in terms of a concrete design dilemma: how can we have students 'in the flow' while gaming within the game-environment as well as active as critical thinkers on a theoretical level transcending the game-environment. This seems to connect to a fundamental design dilemma in inquiry structured or open education aiming at conceptual development described by various authors such as Kirschner, Sweller and Clark (2006) and - for the case of education on electrical circuits - Kock et al. (2013).

Once attained, theoretical insight opens up an avenue to develop towards proficiency with respect to real-life situations and beyond.

2.3 Learning about Electrical Circuits

Learning problems in electricity have been widely documented: Over the years remedies have been suggested to overcome students' conceptual problems in electricity, but only with limited success (Mulhall et al. 2001). The topic is still receiving attention (for example Engelhardt and Beichner 2004; Hart 2008; Taber et al. 2006; Jaakkola, Nurmi & Veermans, 2010). Coming to grips with the scientific concepts in electricity requires an understanding of the physics involved, which is at least partly at odds with the everyday experiences and ways of speaking about electricity (Shipstone 1985; Duit and Schecker, 2007).

A key problem is the development of inadequate conceptual understanding of various aspects of electrical circuits and particular persistent 'misconceptions' that students tend to develop. In Duit's STCSE bibliography on students' 'misconceptions' and conceptual change (Duit, 2009) several hundreds of publications are listed on learning electricity.

Taconis (2010) has described a hierarchical building of concepts concerning electrical circuits

based on a review of literature. Each addition to the concepts requires all concepts in the lower floors to be rightly understood.

1) Correct understanding of electrical circuits essentially being closed but not short circuited, and the electricity circling in the circuit, and the current not being consumed in the circuit,
2) Understanding that two distinct physical quantities are necessary to understand/describe the flow in electrical circuits (and such systems): electrical current and voltage,
3) Understanding of the topological types of electrical circuits; series and parallel and their implications for electrical current and voltage.
4) Understanding of particular electrical components and their properties.

Students may experience problems on either level of this hierarchy and students' alternative ideas often do not correspond to the scientific view and do not easily change through instruction (Shipstone 1985; Duit and Von Rhoeneck 1998; Engelhardt and Beichner 2004; Taber et al. 2006). Kock et al. (2013) conclude: when trying to solve problems or explain phenomena in circuits, students frequently (a) confuse important concepts such as current and voltage, (b) use the idea that current is consumed (or use unipolar, clashing or shared current models), (c) view power supplies as a source of constant current instead of constant potential difference, (d) have difficulties building and drawing circuits and (e) do not realize that a change of one element can have an impact on the current in the whole circuit.

A main obstacle here is that students may tend to understand electrical phenomena in terms of the so called 'experiential gestalt of causation' (Anderson, 1986). This basic misinterpretation may underlay many of the observed misconceptions.
In the 'experiential gestalt of causation' there is an *aim*, a *cause* or chain of causes that instigates a process, a *medium/vehicle*, and a *desired effect*. This mental model implies a number of intuitive qualitative rules such as:

- the effect is roughly in the direction of the cause / chain of causes,
- the stronger the initial cause / chain of causes, the stronger the effect – by default proportional,
- the cause / chain of causes costs 'effort', and is weakened in the long run (due to exhaustion) while the effect continues,
- there is a physical connection of the cause to the effect, possibly through the medium (or vehicle), which may damp the effect – by default proportional to its dimensions,
- the better the medium and / or the smaller the distance the stronger the effect,
- as the cause stops, of if the contact is ended or the medium is removed, the effect stops.

Figure 3 shows an example.

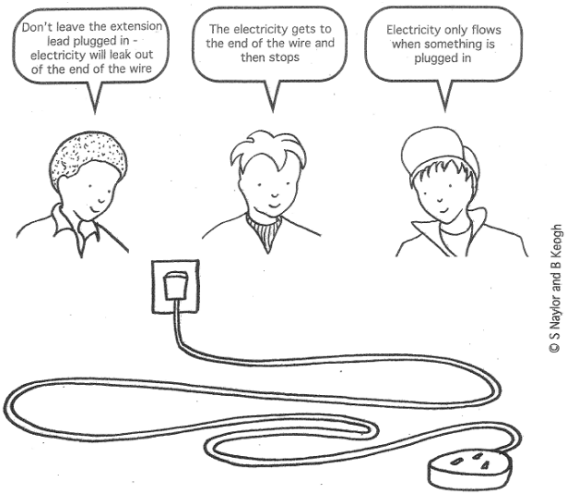


Figure 3: 'Electricity' effect as experienced in day to day life: when plugged in (cause) de 'electricity' from the socket is directed to the light bulb connected (aim) via the cord (medium/vehicle) to produce the desired effect. An example of cause - effect reasoning, from which students may derive interpretations such as: a wire twice as long will make the light bulb half as bright.

An attempt to counter the misinterpretation of electrical circuits from such a linear causal perspective often made is to explain that the electrical circuit is to be understood in terms of an analogy. Two such analogies are regularly used in science education (Hart, 2008) with their own strengths and weaknesses:

a) Fluid current analogies, that of the home heating system in particular,
b) Microscopic analogies in which the electrical current is modelled by a stream of electrons depicted as e.g. lorries carrying an electrical load travelling a closed path.

The E&E electrical endeavours game uses the water current analogy.

2.4 Research Questions

The research questions are:

1. Can we built and use for education a serious game that leads to adequate qualitative understanding of electrical circuits without particular misconceptions?
2. What characteristics of the game and the way it

is used in classroom facilitate adequate understanding of electrical circuits? In particular:
a. in dealing with misconceptions
b. keep a productive the balance between flow-based gaming and reflective and theoretical activities.

3 METHODS

The project was performed in two consecutive rounds each comprising a (re)design and a testing phase. It started with the game version 1.0 as previously build. On advice of the secondary school teachers version 1.0 was upgraded to version 1.1 before starting the project, in order to remove mistakes and smaller difficulties. Version 1.1 was tested by panels of experts in the pedagogy of science, school teachers and students and found adequate for classroom use.

3.1 Description of the Development Rounds

Version 1.1 was taken to the classroom in two Dutch grade 9 classes, one general secondary education and one pre university education. A third one, also pre university education, served as a control group. The lessons were structured to an open inquiry model. The students were allowed to play with version 1.1 during 1½ lesson (approximately 75 minutes). The lessons for the experimental and control groups were partly the same, but deviated in the experimental group playing the game and the control book following the textbook. This is shown in Table 1. Figure 4 a typical screen shot.

In round II, version 3.0 was taken to the classroom in three Dutch grade 9 classes, one general secondary education and two pre university

Table 1: The quality of the circuits built.

	Experimental groups	Control group
1	Introduction Activation of basic knowledge from grade 8	
	Inquiry learning playing with version 1.1 The teacher coaches; no additional teacher explanation of theory	Individual working from the textbook The teacher coaches; no additional teacher explanation of theory
2	Learning task: Build a parallel circuit of two light bulbs	
	Continuation like 1st lesson	Continuation like 1st lesson
3-6	Classical teaching using the textbook	
	Test	



Figure 4: Characteristic screen shot of E&E electrical endeavours version 1.1.

education. Version 3.0 has a renewed level structure that closely follows the regular Dutch grade 9 lesson plan. Also the layout of the game screen is renewed. First, in the left part of the screen the electrical circuit is continuously shown as a closed circuit providing an overview to the students which is much akin to the electrical schemes' usually found in textbooks. Second the working screen on the right is redesigned to more realistically depict voltage and current meters and the way they should be connected. Figure 5 gives an impression.



Figure 5: A characteristic screen shot of E&E electrical endeavours version 3.0.

In round II there was no control group. The lessons were structured according to the teacher guide that was written on the basis of the experiences in the first round and a review of literature. In the first lesson –again- the subject was introduced and the prior knowledge of the students was activated The students were then allowed to play freely with the game for approximately 40 minutes, in which they could take op the challenges built into the games level structure. The next lesson, a classroom

discussion was organized evaluating the students various ideas and results, and drawing common conclusions on understanding electrical circuits. In both coaching the students and the classroom discussions, the teacher could make use of a set of questions in the teacher guide especially designed to stimulate the verbalization/construction of operational rules and theoretical knowledge and to stimulate transfer (see Figure 2). The whole experiment comprised 6 lessons (300 minutes).

3.2 Analysis of the Circuit Building Test

For both rounds, the learning results were obtained using the standard test results (grades). The same test was completed by all groups in that round. In addition an analysis was made of pictures taken from the circuits built in the learning tasks 'build a parallel circuit of two light bulbs'. The analysis focussed on the correctness of the circuits and the presence of misconceptions. Four misconceptions were focussed on in particular: circuit not closed /incorrect, short circuit, a circuit lay-out depicting a ´linear causal 'understanding' of electrical circuits, and the type of meaning of ´parallel´ the students displayed in the circuit built. Concerning the latter: components may be placed 'visually parallel' though when analysing the circuit net ´electrically parallel´ – see figure 6.

A codebook was used to underpin the various judgements on the circuits and the apparent presence of misconceptions and two researchers cooperated in categorizing the various circuits.

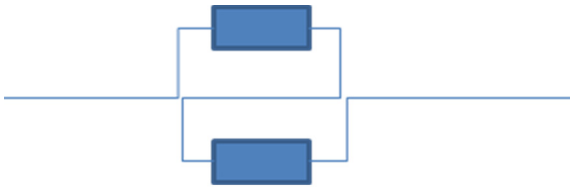


Figure 6: Optically parallel though not electrically.

4 RESULTS

In round I, the teacher reported a marked distinction between students that are experienced gamers and other students. The non-gamers experience difficulty in finding their way in the game, while the gamers seemingly effortless solve the problems presented.
The teacher also reported that all students had difficulty recognizing the electrical components in the game. Students in the experimental setting comment: *"it was fun, but I would have preferred an ordinary lesson since then I would have understood it much better"*.

The result of the analysis of the 'parallel circuits' built in the second lesson and their quality in general is shown in Table 2.

Table 2: The quality of the circuits built.

	Round I			Round II			t-test
	n*	**M**	**SD**	**n***	**M**	**SD**	
g.s.e. exp.	13	-0,54	0,75	11	1,68	1,03	5,94***
g.s.e. contr.	---	---	---	---	---	---	
p.u.e exp.	11	1,05	1,59	24	1,92	0,70	1,74*
p.u.e contr.	12	2,75	0,45	---	---	---	
TOTAL	**36**	**1,04**	**1,70**	**35**	**1,84**	**0,81**	**2.54****

Abbreviations: g.s.e. = general secondary education, p.u.e = pre university education. Sign. Levels: * = 5%, ** = 1%, *** = 0,1%
* duo's of students

The control group that worked from the textbook clearly outperformed the two experimental groups. However it was observed that the average grades for the test at the end of the lessons did not significantly differ for the three groups. The teacher commented that she had to make an extra effort in the experimental groups to secure their progress.

An analysis of misconceptions in the circuits built after the second lesson is shown in Table 3, which shows that misconceptions occurred frequently in the experimental groups.

Table 3: The occurrence of misconceptions in the first and second round.

		Circuit not closed /incorrect	**Short circuit**	**Causal mental model**	**Visual view of 'parallel'**
Round I	M	0,45	0,39	0,52	1,19
(n*=24)	SD	0,47	0,50	0,45	0,38
Round II	M	0,03	0,14	0,41	0,16
(n*=35)	SD	0,17	0,36	0,35	0,81
t-test		-4,19***	-2.10*	-1.0 (ns)	-6,55***

Sign. Levels: * = 5%, ** = 1%, *** = 0,1%
* duo's of students

A qualitative analysis of the circuits built revealed a very interesting pattern.

It appeared that a majority of students (54%) built a circuit in a very particular manner. For general secondary education students this is even 77%. First they added a switch to the circuit - a component that was not at all mentioned in the assignment. Second they ordered the components in a linear fashion starting with the switch, followed by the 'parallel part' and ending with a light bulb. Figure 7 shows three examples (three left columns),

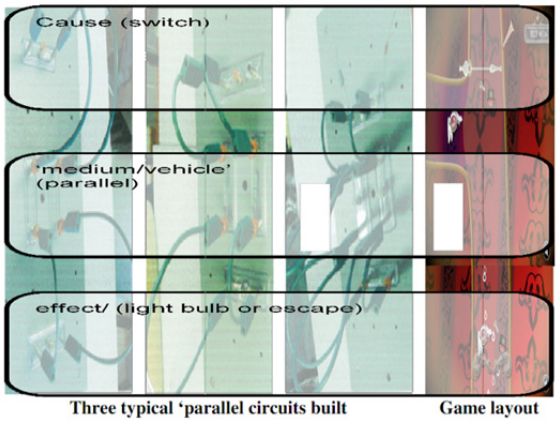


Figure 7: Student mimicking the games circuit lay-out and linear topology when building a parallel circuit.

as well as the game lay-out (right column). This typical lay-out by the students both resembles the game's screen lay-out and is in accordance with the 'experiential gestalt of causation'.

The disappointing learning result from the game as such, the lack of student and teacher enthusiasm, the frequent occurrence of misconceptions, and the apparent strong but unwanted learning effect arising from the games lay-out that effectively *encourages* students to understand electrical circuits in terms of 'the experiential gestalt of causation' formed a strong incentive to improve both the game, and the way it is used in classroom. First a panel of teachers was asked to comment on the game and the inquiry based classroom implementation. This resulted in redesigning specifications for the game, and the conclusion that the open-inquiry approach should be replaced by a carefully planned approach combining experiential learning on one hand and reflection, verbalisation en additional experience on the other. This criticism was theoretically underpinned by experimental work of Kock et al. (2013) and the theoretical insight by Kirschner, Sweller and Clark (2006). A review of literature led to the model by Taconis (2011) depicted in Figure 2, and a extend teacher guide comprising: a) background information on learning from serious games and understanding electrical circuits, b) an example time table, c) concrete questions the teachers could use to stimulate the verbalization/construction of operational rules, theoretical knowledge and their transfer (Figure 2).

In round II the changes were implemented. The teachers reported the students being enthusiastic about the game, and students reporting that 'it really helped them to understand electrical circuits'. One student for instance stated: *'the game really made me understand what I am doing, and helps me to explain what it is to my friends'*. This was supported by a very high mean score on the test concluding the lessons: 7.2 out of 10 (usually 6.3).

The quality of the 'parallel circuits' built is shown in Table 2. It reveals a significant increase with respect to the quality of the circuits built in the first round. Table 3 shows that also the occurrence of misconceptions significantly decreased, except for the occurrence of the 'experiential gestalt of causation'.

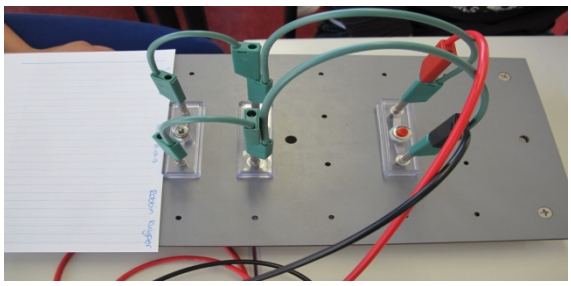


Figure 8: Typical wrong placement of the switch.

A qualitative analysis of the circuits built by the students revealed that in addition many students (54%) have difficulties with the correct placement of the switch (Figure 8). It effectively functions as an optional 'shortcut' directly at the entrance of the powering wires from the power unit. Again, this is in clear accordance with de games lay-out (Figure 5), were the switch has taken the form of a beam across and blocking the current at the top of the screen.

4 CONCLUSIONS & DISCUSSION

We conclude that the close cooperation between school teachers, experts on science education and game-designers was a successful way to a clear improvement in both the game and the way it is adequately used in classroom. Note that enhancing the games contributed to cognitive learning did not imply a decrease in student enthusiasm. The progress made is documented in the projects physical products: the game and guideline for its adequate use in education.

It is concluded that the game when used according to the guideline, probably contributes to improved student understanding. However, misconceptions still occur, those related to 'the experiential gestalt of causation' in particular. In terms of the model by Taconis (2011) learning effects seem to concentrate on the level of 'skilled action' (building an electrical circuit) and

´operational rules´. Students do not report on theoretical understanding or models they (re-)constructed.

Concerning the second research question, it is found that the games screen lay-out has a strong though undesired impact on students' mental model of electrical circuits and student learning. The way things are presented apparently strongly influences the students´ way of 'looking at things'. This 'topological mimicking' seems to be a very powerful learning mechanism. The games lay-out, and probably its structure as well, sends out a very strong message about 'how things are'. Though not entirely effective in the case of our project, the strength of this mechanism is potentially valuable for designing serious games.

Concerning the use of serious games in classroom, it became clear that 'open inquiry' alone is not the way to go. An approach balancing gaming and deep cognitive processing in alternating phases seems much more fruitful. In this a rule of thumb helping teacher in a practical way could be: "treat serious games as experiments or practical exercises".

Our research has been on a fairly small scale, and as such its outcomes may be easily over generalized. But they point toward two particular issues of importance for larger scale future research:

- developing a pedagogical approaches balancing between play and open inquiry on one hand and stimulating and structuring deep cognitive processing on the other. Put more generally: seeking a productive balance between flow and reflection.
- taking advantage of the strong modelling impact the lay-out and the structure of the game have on student learning.

ACKNOWLEDGEMENTS

This research was funded by the SLOA fund of the Dutch 'onderwijs cooperation' seeking to promote secondary school teachers doing educational research in cooperation with academic educational researchers.

REFERENCES

Anderson, B. (1986). The experiential gestalt of causation: a common core to pupils' preconceptions in science. *European journal of science education*, *8*, 155-171.

Csikszentmihaly, M. (2009). *Creativity: Flow and the Psychology of Discovery and Invention*. HarperCollins, New York.

Csikszentmihalyi, M.(2000). *Das Flow-Erlebnis. Jenseits von Angst und Langeweile: im Tun aufgehen* [Beyond boredom and anxiety] (8th. Ed.).Stuttgart, Germany: Klett-Cotta.

Dekker, J., Jacobs, A., van Hensberg, V. (2010). *Informal user test of E&E electrical endeavours*. Internal report.

Dreyfus, S., E., & Dreyfus, H. L. (February 1980). A Five-Stage Model of the Mental Activities Involved. In *Directed Skill Acquisition*. Washington, DC: Storming Media. Retrieved June 13, 2010. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA084551&Location=U2&doc=GetTRDoc.pdf.

Duit, R. (2009). *Bibliography STCSE: Students' and teachers' conceptions and science education.* http://www.ipn.uni-kiel.de/aktuell/stcse/stcse. html. Accessed 20 April 2009.

Duit, R., & Schecker, H. (2007). Teaching Physics. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 599–629). New York: Routledge.

Duit, R., & Von Rhoeneck, C. (1998). Learning and understanding key concepts of electricity. *In A.Tiberghien, E. Leonard Jossem, & J. Barojas (Eds.), Connecting research in physics education with teacher education*. International Commission on Physics Education.

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, *72*(1), 98–115.

Gravemeijer, K., & Cobb, P. (2006). Design Research from a Learning Design Perspective. In J. Van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research*. London: Routledge.

Hart, C. (2008). Models in physics, models for physics learning, and why the distinction may matter in the case of electric circuits. *Research in Science Education*, *38*, 529–544.

Jaakkola, T., Nurmi, S., & Veermans, K. (2010). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. *Journal of research in science teaching*. http://onlinelibrary.wiley.com/doi/10.1002/tea.20386/pdf.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, *41*(2), 75-86. http://igitur-archive.library.uu.nl/fss/2006-1214 - 211848/kirschner_06_minimal _guidance.pdf.

Kock, Z. D. Q. P., Taconis, R., Bolhuis, S. M & Gravemeijer, K.P.E. (2013). Some key Issues in creating inquiry-based instructional practices that aim at the understanding of simple electric circuits. *Research in Science Education*, *43*(2), 579-597.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. New Jersey: Prentice Hall.

Koops, M. C. & Hoevenaar, M. (2012). Conceptual Change During a Serious Game: Using a Lemniscate Model to Compare Strategies in a Physics Game. *Simulation & Gaming.* http://sag.sagepub.com/content/early/2012/09/23/1046878112459261.abstract.

Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train and inform.* Thomson: Boston, MA.

Mulhall, P., McKittrick, B., & Gunstone, R. (2001). A perspective on the resolution of confusions in the teaching of electricity. *Research in Science Education*, *31*, 575–587.

Ruben, B. D. (1999). Simulations, games, and experience-based learning: The quest for a new paradigm for teaching and learning. *Simulation & Gaming*, *30*(4), 498–505.

Shipstone, D. (1985). Electricity in Simple Circuits. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children' s ideas in science*. Milton Keynes: Open University Press.

Taber, K. S., de Trafford, T., & Quail, T. (2006). Conceptual resources for constructing the concepts of electricity: the role of models, analogies and imagination. *Physics Education*, *41*(2), 155–160.

Vockell, E.L. (2013). *Educational Psychology: A Practical Approach*. Chapter 6. Retreived from Perdue University sept. 2013http://education. purduecal. edu/Vockell/EdPsyBook/Edpsy6/edpsy6_transfer.htm.

APPENDIX

The E&E game and its versions:
1.0 and 1.1 (http://www.eeee.tue.nl/)
2.0 (http://www.eeee.tue.nl/sloa/index.html)
3.0 (http://www.eeee.tue.nl/sloa/thebeginning/index.html)

Automated Generation of Geometry Questions for High School Mathematics

Rahul Singhal[1], Martin Henz[1] and Kevin McGee[2]
[1]*School of Computing, National University of Singapore, Singapore, Singapore*
[2]*Department of Communications and New Media, National University of Singapore, Singapore, Singapore*
{*rahulsinghal, mckevin*}*@nus.edu.sg, henz@comp.nus.edu.sg*

Keywords: Automated Deduction, Graph-based Knowledge Representation.

Abstract: We describe a framework that combines a combinatorial approach, pattern matching and automated deduction to generate and solve geometry problems for high school mathematics. Such a system would help teachers to quickly generate large numbers of questions on a geometry topic. Students can explore and revise specific topics covered in classes and textbooks based on generated questions. The system can act as a personalized instructor - it can generate problems that meet users specific weaknesses. This system may also help standardize tests such as GMAT and SAT. Our novel methodology uses (i) a combinatorial approach for generating geometric figures (ii) a pattern matching approach for generating questions and (iii) automated deduction to generate new questions and solutions. By combining these methods, we are able to generate questions involving finding or proving relationships between geometric objects based on a specification of the geometry objects, concepts and theorems to be covered by the questions. Experimental results show that a large number of questions can be generated in a short time. We have tested our generated questions on an existing geometry question solving software JGEX, verifying the validity of the generated questions.

1 INTRODUCTION

Geometry, the study of space and spatial relationships, is an important and essential branch of the mathematics curriculum at all grade levels. The study of geometry develops logical reasoning and deductive thinking, which helps us expand both mentally and mathematically. Children who develop a strong sense of spatial relationships and who master the concepts and language of geometry are better prepared to learn number and measurement ideas, as well as other advanced mathematical topics (National Council of Teachers of Mathematics, 1989).

Euclidean geometry is a branch of mathematics which deals with the study of plane and solid figures on the basis of axioms and theorems employed by the Greek mathematician Euclid. It is important to understand Euclidean geometry when studying a course because geometry does not follow any set pattern. In Euclidean geometry, one can only learn the axioms and results proven from these axioms. The student must apply these axioms with no set pattern or list of steps for solving such questions. Therefore, a question may have (possibly infinitely) many solutions. To practice the required problem solving skills, students require a large number of different types of geometry questions on various concepts. Generally, textbooks and online sites provide a limited predefined number of questions for each topic. Once practiced, these questions lose their purpose of enhancing student thinking. The tedious and error-prone task of generating high-quality questions challenges the resources of teachers. Hence, there is a need for software which assists both teachers and students to generate geometry questions and solutions.

The software can also act as a personalized instructor. It can generate questions that cover the required topic and meet the required level of student proficiency. Apart from helping users, the framework of generating questions has scientific contributions to other research areas, such as Intelligent tutor systems (ITS) and Massive Online Open Courses (MOOC).

Various research has been performed in automated deduction of theorems at high school level in the geometry domain, although none with the goal of automatic question generation. Instead, they mainly demand users to generate the question with the help of tools. In addition, they mainly focus on solving and assessment of the questions. Our survey shows that the currently available geometry systems, such as

JGEX, Geogebra, Cinderella and Sketchpad, are not able to automatically generate questions of user specified geometry topics.

The aim of this paper is to develop a framework that can be used to generate geometry questions based on specific inputs, such as geometry objects and theorems to be involved in their solution. For a given set of geometry objects, the algorithm can generate a large set of questions along with their solutions. The solutions will involve user desired theorems directly or indirectly. Hence the framework can generate questions to test the theorem on various geometry objects and concepts.

Our framework can generate questions involving algebraic computations for a solution. Currently, we restrict the relationships between quantitative entities to linear ones. Our framework has a predefined database of theorems and concepts which can be used for generating questions. Given a set of user-selected objects, our system can generate all possible questions using an existing database of concepts and theorems.

The main contributions of this paper are as follows:

1. Our geometry question generator combines the complementary strengths of a combinatorial approach, pattern matching and deductive reasoning. It can generate geometry questions which were not possible previously.
2. A substantial evaluation is provided that demonstrates the effectiveness of our generator. It can generate various categories of the questions covered in the textbooks and questions asked in SAT and GMAT.
3. A knowledge representation is described for geometry objects and predefined theorems. This representation helps in applying theorem information within the generated questions.

2 RELATED WORK

In this section, we provide a general review of related works. Computational research in the geometry domain started in the 19th century. However, lack of question generating in geometry in the literature required a principled ab-initio approach in our work.

Researchers mainly focused on proving geometry theorems. Broadly, geometric theorems are proven using algebraic or non-algebraic methods. Algebraic methods such as Wu's method and Gröbner bases (GB) (Chou et al., 1994) generate algebraic equations from the given facts and relations. They involve coordinate geometry and algebraic formulas to find new theorems. Hence, the proofs generated by these methods are out of scope of high school mathematics (Chou and Gao, 2001). Therefore, we are mainly interested in the non-algebraic methods. Non-algebraic methods for automatic discovery and proof of geometry theorems can be further divided into three approaches: coordinate-free methods, formal logic methods and search methods.

2.1 Coordinate-free Methods

These methods are applicable to constructive geometry statements of equality. Various methods have been proposed under this category, such as the area method (Narboux, 2010; Chou et al., 2011), the full-angle method (Chou et al., 2000; Wilson and Fleuriot, 2005), the complex number method, the vector method for Euclid plane geometry (Chou et al., 2000), the volume method for Euclidean solid geometry (Chou and Gao, 2001) and the argument method for non-Euclidean geometry (Chou and Gao, 2001). The area method was further improved and developed into a computerized algorithm (Chou et al., 2000; Chou and Gao, 2001). These methods can only be applied in constructive geometry, which is outside the scope of this paper.

2.2 Formal Logic Methods

Theorems in Tarski's geometry were proven using Interactive Theorem Prover (ITP) (McCharen et al., 1976), albeit limited to several trivial theorems. ITP is an interactive theorem prover based on the resolution principle, which generates resolution style proofs that resemble traditional proofs. In 1989, Quaife continued the work of McCharen with Otter (Quaife, 1989). Otter is an automated theorem prover based on resolution. A series of tactics such as Hyper-resolution, UR-resolution, paramodulation, support set and clause weight can improve the resolution efficiency. In 2003, Meikle and Fleuriot developed Hilbert's geometry with the theorem prover Isabelle/Isar (Meikle and Fleuriot, 2003), an interactive and/or semi-automated theorem prover. The greatest disadvantage of formal logic methods is their low reasoning efficiency (Jiang and Zhang, 2012), caused by combinatorial explosion of their search space.

2.3 Search Methods

Fundamentally, the search method is used in a rule-based expert system, which includes a rule database,

a fact database and a reasoning engine. The inference rules stored in the rule database include axioms, theorems, lemmas, formula, definitions and algebraic operation rules in geometry. Geometric facts stored in the fact database include geometry predicates such as angle bisector, equidistant points, parallel lines and perpendicular lines. The reasoning engine deduces new geometric facts by applying inference rules to the facts database. There are three ways of performing deduction search, namely forward chaining (Zhang et al., 1996; Wilson and Fleuriot, 2005), backward chaining (Wilson and Fleuriot, 2005), and bidirectional chaining (Coelho and Pereira, 1986).

Forward chaining starts from the hypotheses and rules and works towards the conclusion, while backward chaining starts from the conclusion and works towards the hypotheses through rules, and the bidirectional chaining proceeds in both direction simultaneously. In 1975, Nevins proved geometry theorems via bidirectional chaining (Nevins, 1975), and in 1986, Coelho and Pereira developed a prover GEOM based on bidirectional chaining (Coelho and Pereira, 1986). Unlike Nevins, GEOM implemented backward chaining in the reasoning process and only used forward chaining to search congruent triangles hidden in the diagram. In 1995, Chou, Gao, and Zhang described an integration of a deductive database into the search methods (Chou et al., 2000; Gao et al., 1998; Gao et al., 1998; Chou and Gao, 2001). The deductive database method can find a fixpoint for a given geometry diagram, containing all properties of the geometry diagrams that can be deduced using a fixed set of geometry rules. They effectively controlled the size of the facts database with structural deductive database techniques.

Each search method has different advantages and disadvantages. Forward chaining is always feasible, but does not have an explicit reasoning goal. Backward chaining has an explicit reasoning goal, but sometimes lacks feasibility. The bidirectional chaining method is feasible and has an explicit reasoning goal, but is hard to implement. We are using forward chaining in our framework as it is most suited for generating previously unknown quantities.

3 GEOMETRY QUESTION SPECIFICATION

Mathematically a geometry question Q generated by our system can be represented by a quintuple (Object O, Concept C, Theorem T, Relationship R, Query type qt) where:

- O ∈ (lines, triangles, square, circle, ...)

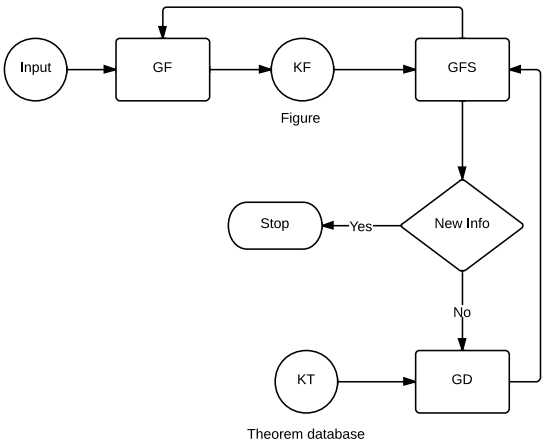


Figure 1: Connection of these components and knowledge representations.

- C ∈ (perpendicular, parallel, midpoint, angle-bisector, circumcircle...)
- T ∈ (Pythagorean theorem, similarity theorem, various triangle-theorems, ...)
- R ∈ (syntactic, quantitative)
- qt ∈ (syntactic, quantitative)

In order to generate geometry questions, the user has to provide a set of geometry objects O such as triangles, squares, etc., and a set of concepts C which the user wants the generated question to cover. Optionally the user may select a set of theorems T to be tested by the question. The relationship R can be either syntactic such as perpendicular, parallel, etc., or quantitative such as the length of an object, the ratio of two quantities etc. The query type qt is the type of generated question that can be asked to find the hidden relationship which can be calculated from the given information.

4 FRAMEWORK

Our framework comprises three major components along with the knowledge databases used for storing input, geometry figures and a set of predefined theorems. Figure 1 shows the connection of these components. The input consists of geometric objects, concepts and theorems selected by the user. The input is fed into the first component, *Generating Figure (GF)*. This component is used for generation of geometric figures from the input. Each figure constitutes a diagrammatic schema (DS) (Greeno et al., 1979) and a set of unknown variables representing the relationship between geometric objects. The geometric figure is passed to the second component, *Generating*

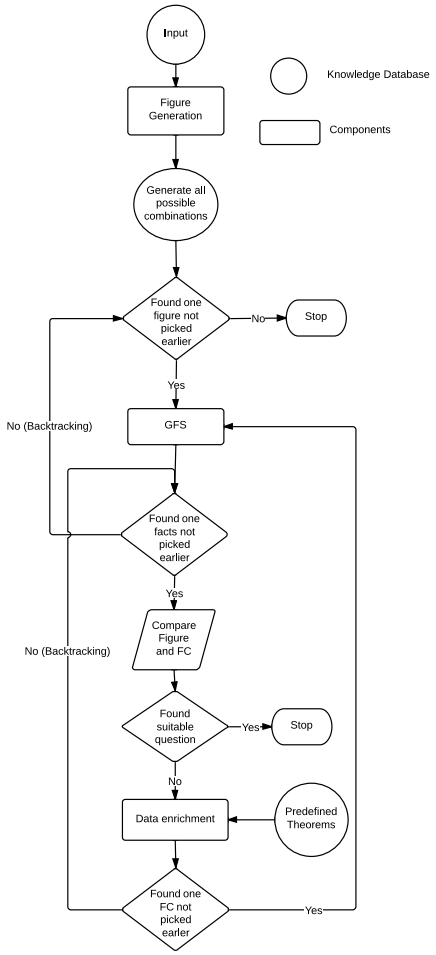


Figure 2: Flow diagram of the algorithm for generating questions.

Facts and Solutions (GFS). This component is used to find values of the unknown variables representing possible relationships to be asked by the generated question. GFS makes use of the predefined knowledge database of axioms. It results in the formation of a *configuration* (Cfg) containing known values for some relationships between its objects. New information refers to the generation of suitable questions. A question is considered suitable if it covers the essential information such as a new fact and a proper reasoning for the generated fact. Currently, the decision of suitability is taken manually by the user.If the suitability conditions for the generated configuration (Cfg) are not met then configuration is fed into the last component, *Generating data for the figure (GD)*. GD assigns values to unknown variables of relationships. Repeated processing by GD makes the questions generated from Cfg easier and easier, because the values assigned by GD appear as given facts in the generated questions. GD makes use of a predefined set of theorems and makes sure that the assignment results in successful generation of geometry questions. FC generated from this component is again passed to GFS component and this loop continues until a question is found which meets suitability condition.

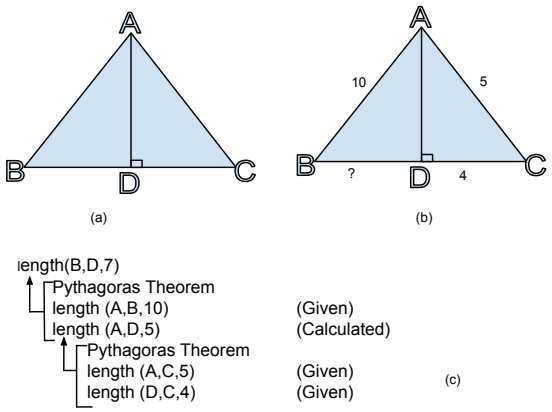


Figure 3: (a) The figure is generated by using *triangle* and *perpendicular* as the geometry objects using the GF component. (b) The data is generated for (a) using the GD component. (c) The new fact and its derivation using the GFS component is generated from the figure shown in (b).

Along with these components, two knowledge databases are used in the framework, namely the Knowledge database for the generated figures *(KF)* and the knowledge database for predefined theorems *(KT)*. KF is generated dynamically during generation of questions while KT is fixed and can be modified only after the algorithm terminates. KF is generated and modified by the GF function, and GD requires both KF and KT for generating data for the figure. GFS has a set of axioms but no access to any knowledge databases.

4.1 Algorithm

Figure 2 represents the flowchart describing the algorithm for generating geometric questions. Algorithm 1 describes the flowchart in further detail. Here, we shall explain Algorithm 1 with the help of an example. Figure 3 shows the step by step execution of the algorithm. We select the following input in our example.

- Object: triangle and line segment
- Concept: perpendicular
- Theorem: Pythagorean Theorem
- Number of questions: 1

In the next subsections, we will describe each knowledge database in detail, followed by the three components and their interaction with these databases.

Data: User selects object(s), concept(s), theorem(s) and the number of questions to be generated.
Result: Question with single or multiple solutions.

1. Generate all possible figures consisting of geometry objects using GF function from the given input.
2. Find one figure which has not been picked earlier. If found, go to next step else terminate.
3. Save this figure using KF knowledge database.
4. Assign values to variables of figure obtained in second step from the predefined knowledge database of axioms through GFS function. Configurations (Cfg) are generated from this step.
5. Find one Cfg not picked earlier. If found, goto next step. If not found, backtracking to step 2.
6. Compare Cfg with the previously stored figure.
7. If the comparison gives the desired suitability then the Cfg is declared as a generated question and the algorithm stops. If the conditions are not met or more number of questions are required, go to the next step.
8. Configuration Cfg obtained from GFS is fed into the third component to assign more unknown variables. It makes use of KT, a predefined database of theorems.
9. From the configurations obtained from GD, find a new configuration Cfg which has not been chosen earlier. If found, goto step 3. If all configurations have been chosen earlier, goto step 5.

Algorithm 1: Algorithm for generating geometry questions.

4.2 Knowledge Database for the Figure/ Figure Configuration (KF)

KF contains the question figure and configurations using a graph-based knowledge representation. The

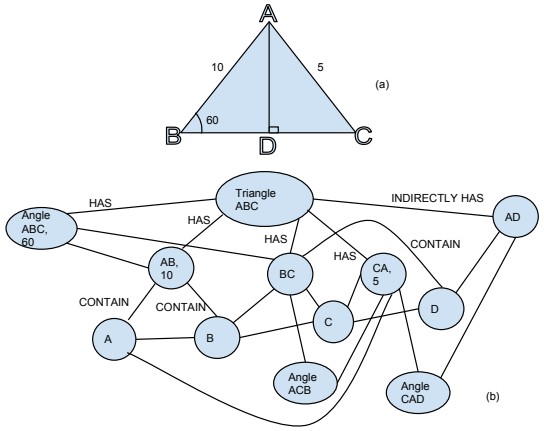


Figure 4: (a)The geometry figure and (b) Partially drawn knowledge representation of (a).

nodes of our graph represent the geometric objects. Two nodes can have multiple labeled arcs between them, representing multiple relationships. In addition, an arc can be bidirectional or unidirectional depending on the relationship between the nodes. Following the example of the previous section, Figure 4(a) shows the geometric figure: a triangle ABC is given and AD is perpendicular to BC. Figure 4b shows its partial representation in a graph format where we focus on the most important relations. Relationships such as "equal angle" and "equal length", which require multiple objects/nodes, are represented by nodes, e.g. , the relationship *"∠ABC"* in Figure 4(b). In addition, nodes of our graph store the value of quantitative relationships. For example, the node named *" Angle ABC "* stores the value of angle and is connected to the two nodes representing sides AB and BC.

KF is generated dynamically from the geometry objects selected by the user. KF is initialized by the GF component and later modification is done according to Algorithm 1 defined in Section 4.1. Modification involves creation and deletion of nodes and arcs between them.

4.3 Knowledge Database for Predefined Theorems (KT)

KT is a knowledge database which contains the predefined database of theorems. KT is used for assigning values to the unknown variables of configurations generated by the GF component. KT uses a graph-based representation similar to KF. However the difference lies in the timing of generation. Unlike KF, KT is generated before the system can generate questions and remains unchanged during the algorithm execution. In order to use KT in the algorithm, some

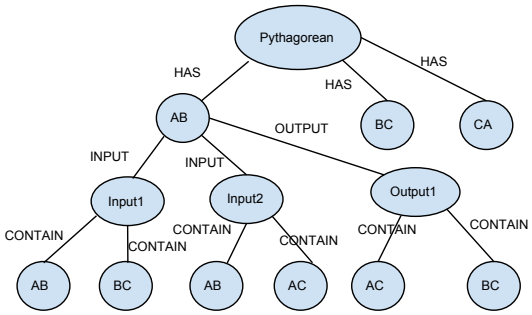


Figure 5: Representation used for generating data for the given figure.

properties of theorems need to be known. Each theorem can be applied to a particular geometry configuration. Each theorem has certain inputs and outputs corresponding to inputs in the geometry configuration. Hence usage of a theorem requires assigning the variables of input and/or output. In KT, the information of input and output for each theorem is stored along with the geometric configuration in which the theorem can be applied. Nodes represent objects whereas arcs represent the input and output relationships. Input and output are decided offline before the execution of the algorithm and later used for assigning unknown variables to get a useful question. To illustrate the process, the Pythagorean Theorem is taken as an example in Figure 5. Given a triangle ABC, where $\angle ABC$ is 90°, Figure 5 partially shows its usage of KT in generating data. The node representing the side AB has three arcs, two of which represent input and one of which represents an output relationship. To use the side AB as an input, the length of the sides AB and BC or AC needs to be assigned. On the other hand, to use side AB as an output, the length of sides AC and BC needs to be given. By giving this as input or output, we can be sure that the Pythagorean Theorem will give a consistent result.

Now we are ready to describe the three components which will use these knowledge databases to generate questions and answers.

4.4 Generating Figure from the User Input (GF)

This is the first step executed by Algorithm 1 described in Section 4.1. This component generates a figure through the combination of a predefined number of ways to combine geometric objects. Currently, we are focusing on triangles and line segments. Hence our algorithm includes combinations in which various triangles and lines can intersect. Furthermore, we are currently limiting our algorithm to the intersection of two triangles. Adding more objects may

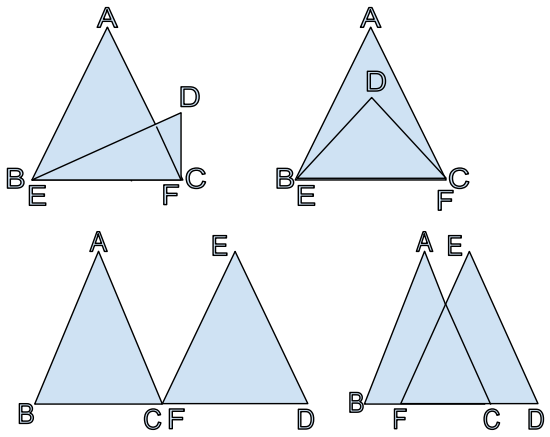


Figure 6: Some predefined ways two triangle can intersect.

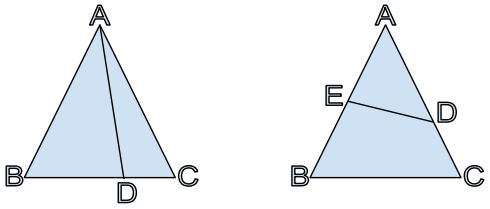


Figure 7: Some predefined ways a triangle and a line can intersect.

make the figure more complex, possibly leading to unusual questions that are not commonly found in textbooks. However there is no limitation on the intersection of lines and triangles. Figure 6 shows some of the possible ways of intersection of two triangles, which have been taken into consideration while generating questions. Intersection of two triangles includes various cases such as two triangles sharing common sides, common part of a side, common vertex. Figure 7 shows the two ways in which a line can intersect with a triangle. Figure 8 describes the flowchart for figure generation from the user-desired input. It includes three functions for adding various items such as objects, concepts and theorems. These functions will add their respective items into the generated DS. For instance, addObject will add a new object without removing the old one. Figure 9 shows an example of adding a new object. Figure 10 shows the same functions but for concept involvement in the DS. Similar functions are used for adding theorems, not discussed here.

The algorithm starts looking for DS in our knowledge database, comprising of the user-desired input. If DS is not found, then a DS is generated using the addObject function, possibly resulting in several DS. From the collection, a DS is picked and moved to the next level of adding concepts to it. If each DS has been picked, our algorithm terminates. The same approach is applied for concept addition in an existing

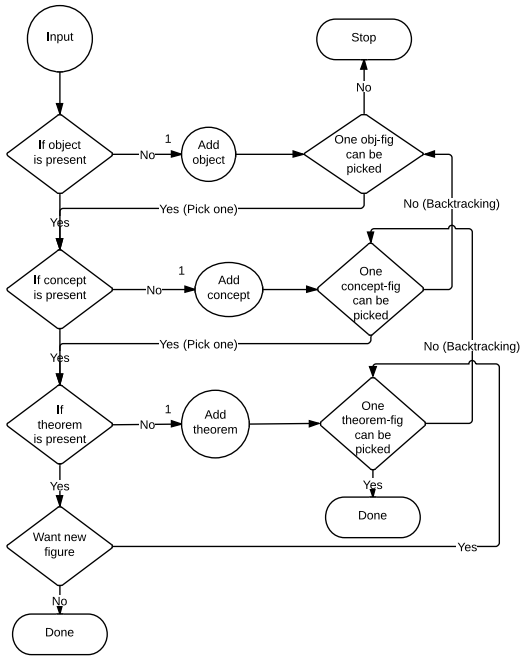


Figure 8: Working of GF component.

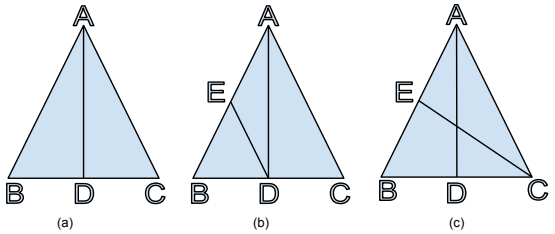


Figure 9: (a) Original figure; (b, c) figures resulting from calling addObject function which adds one more line to the configuration.

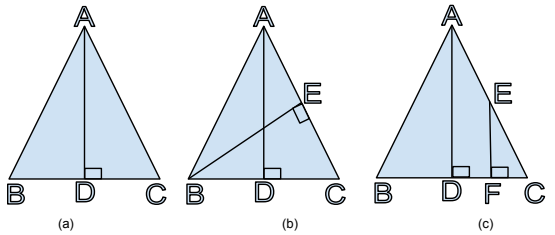


Figure 10: (a) Original figure; (b, c) figures resulting from calling addConcept function which adds one more concept to the configuration.

DS. However, in case all DS formed after concept addition have been picked, backtracking is performed and a new DS is picked from the object-figures. A similar approach is taken when all theorem-figures have been picked, where backtracking leads to choosing a new concept-figure from the previously generated collection.

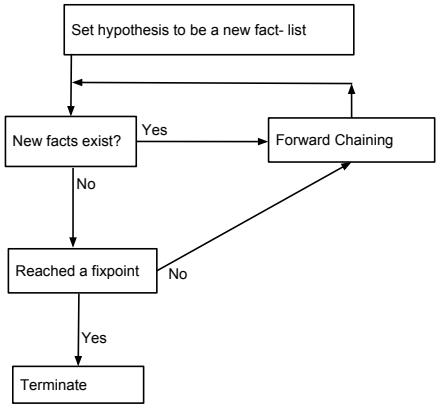


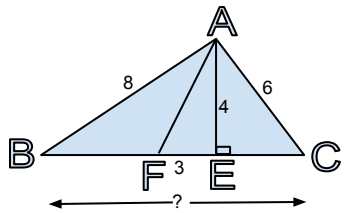
Figure 11: Represents the approach used by GFS component in finding new facts.

4.5 Generating Facts and Solutions from the Configuration (GFS)

This component is responsible for finding the values of unknown variables of the generated figure/configuration from the other two components. This component acts as question generator and solver. The unknown variables whose values have been found represent the generated questions. The steps that leads to finding the unknown variables represent the solution. There can be many ways for finding the values of the unknown variables. In such cases, this component shows all solutions. For generating new facts, it uses a predefined database of theorems. The difference between the theorem database of this component and the previous two components lies in the representation of theorems. In GFS, each theorem is represented in the form of an axiom, while GQ and GF represent theorems in the form of KT and KF, respectively. In addition, GS includes a few basic algebraic theorems, such as $(a+b)^2 = a^2 + b^2 + 2ab$, which are not present in KT and KF. Such algebraic theorems are used inside the axiom system for solving and generating questions which involve algebra.

School Mathematics Study Group (SMSG) axiomatic system is used, which is a combination of Birkhoff's and Hilbert's axiom (Francis, 2002) system. Figure 11 shows the flow diagram of this approach. A fixpoint is a state in which no new facts—in our case, no new assignment of unknown variables—can be deduced from the given facts (Chou et al., 2011). For reaching a fixpoint, forward chaining is used to infer new facts from given facts and rules. The rules used by our system are described in Prolog as follows:

$$Q(x,n_s) : -P_1(x,n_1)...P_k(x,n_k) \qquad (1)$$



(a)

length(B,C,10)
Apollonius Theorem
length (A, B, 8) (Given)
length (A, C, 6) (Given)
length (A, F, 5) (Calculated)
Pythagoras Theorem
length (A, E, 4) (Given)
length (F, E, 3) (Given)

(b)

Figure 12: (a)A question and its solution(b) involving measurement of geometry objects which can be solved with our rules.

which means

$$\forall x[(P_1(x,n_1) \wedge ... \wedge P_k(x,n_k)) \Rightarrow Q(x,n_s)] \quad (2)$$

where $n_1,..,n_k$ and n_s represent numeric values. This representation results in the storage of the integer/real value of the geometry object in last argument of the predicate. With the help of new rules, we can now represent the facts related to dimensions of the geometric objects. Hence this representation of rules can generate geometric questions, which involve measurements of the geometric objects.

In addition, we have added predefined algebraic functions which can be used in rules. One instance of our new rule is shown in Equation 3. The predicate *sq* denotes the square function, which computes the square of a and stores its value in b.

$$sq(a,b) \quad (3)$$

Figure 12 shows an example which can be solved by our rules. Given a triangle ABC, $AE \perp BC$, AB = 8, AC = 6, AE = 4, FE = 3. *our goal is to find the length of BC*. Figure 12(b) provides a solution of this question with the help of our new rules. In addition, these rules help to represent the basic theorems which involve measurement values such as the Pythagorean theorem, the Triangle Inequality theorem and the side/angle ratio theorem in a triangle. Hence it is possible to generate geometric questions based on the above mentioned theorems. For instance, the Pythagorean theorem can be represented in the following way:

$$diff(X,C,0) : -Triangle(a,b,c), sq(a,A), sq(b,B), sq(c,C), sum(A,B,X). \quad (4)$$

A trivial question which requires the Pythagorean theorem for solving would be as follows. *Given a right angle triangle and any two sides of the triangle, the goal is to find the third side.*

Similarly, the Triangle Inequality Theorem can be represented in the following way:

$$lessthan(X,0) : -Triangle(a,b,c), length(a,A), length(b,B), length(c,C), sum(A,B,Z), diff(Z,C,X). \quad (5)$$

A trivial question from this theorem would be as follows. *Given the length of any two sides of a triangle, our goal is to find the maximum value of the third side*.

4.6 Generating New Configurations from an Existing Configuration (GD)

Figure 13 explains the algorithm in the form of a flowchart. It starts with searching for user-selected theorems. In case the user has not chosen any theorem, a theorem is selected from the predefined knowledge database of theorems. However, our algorithm terminates when all theorems have been selected once for a given configuration. After the theorem selection, pattern matching on the theorem figure is performed. If a matching pattern is found, a set of input and/or output values are assigned to the chosen pattern. Assignment is done with the help of a predefined set of theorems. In case the pattern is not found, a new theorem is chosen from KT and the whole process is repeated until we get the desired configuration.

Depending on the requirements, several theorems can be involved in a given configuration. In addition, it is possible to use a single theorem multiple times. The method requires the concept of using the output of one theorem as the input of another one. In this way we can have all possible combinations of predefined theorems. Figure 14 shows an instance of assigning unknown variables. Figure 14b represents the new data assigned to Figure 14a. It is generated with the help of applying Pythagoras Theorem in triangle ABE. The output is changed from BE to AE. Output is considered the length of side calculated using the Pythagoras Theorem from a given input. The input and output for the Pythagoras Theorem is taken using KT. Figure 14c is generated by adding more data in Figure 14a. New data is assigned by applying the

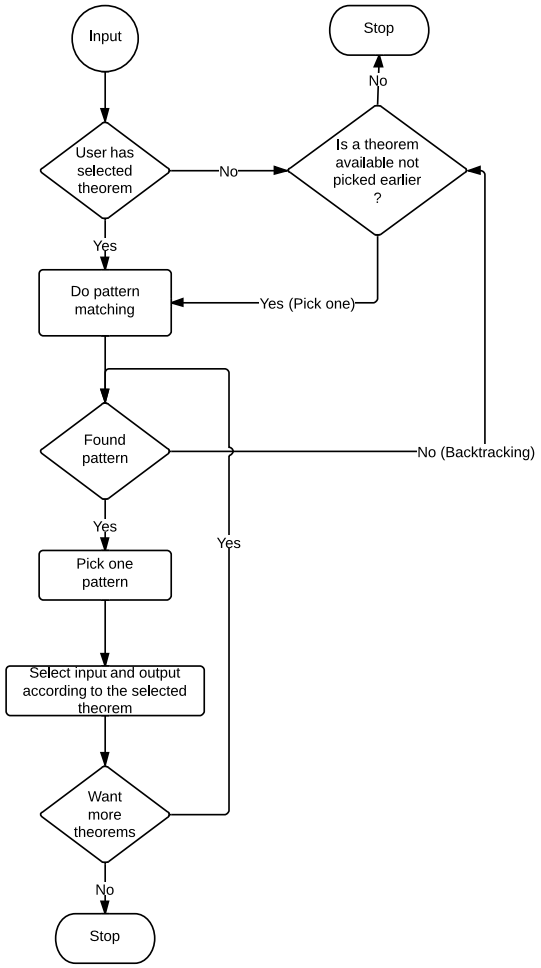


Figure 13: Working of GD algorithm.

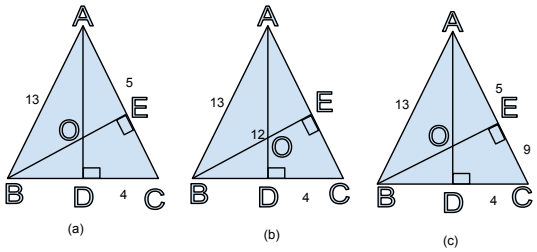


Figure 14: (a) A figure and data is assigned to few geometry objects. (b) Output of calling the changeData function. (c) Result obtained by calling the addData function on (a).

Pythagoras Theorem in the triangle ADC. The theorem requires two sides as input to generate one output. Hence length of side EC is given.

In the above algorithm, it may happen that redundant data is provided in the figure. Redundant data refers to the values assigned to properties of some of the geometric objects which are not required and can be derived from the previously given data and predefined theorems. Figure 15 explains this situation with the help of an example. In this figure, one data item is redundant and any one of the length values can be removed, since, by applying the similarity concept, the third length can be derived from the other two lengths. This type of reasoning may lead to an increase of the difficulty level of the generated questions. This reasoning is added with the help of GFS, which is described in Section 4.5.

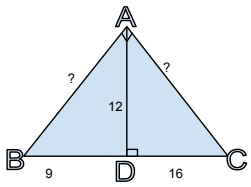


Figure 15: An example in which redundant data is provided in the given figure.

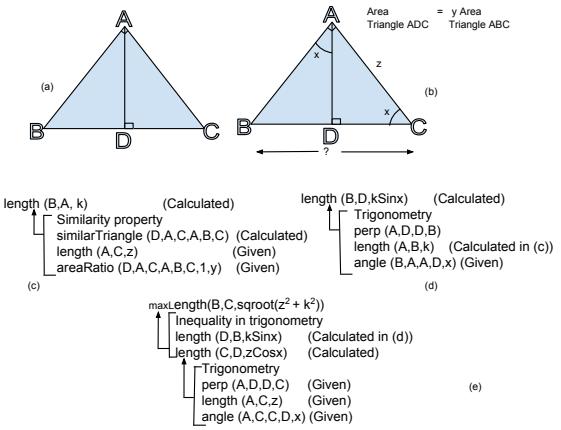


Figure 16: (a) The figure is generated by using *triangle* and *perpendicular* as the geometry objects using the GF component. (b) The data is generated for (a) using the GD component such that the two triangles ADC and ADB become similar. (c, d, e) The new fact and its derivation using the GFS component is generated from the figure shown in (b).

Figure 13 contains non-deterministic decisions at various steps. The non-deterministic selection of options can be avoided by adding more heuristics such as fixing the number and types of theorems to be involved.

5 FURTHER EXAMPLES

This section explains our framework through more examples covering various concepts and geometric theorems.

5.1 First Example

The input given in this question generation task is as follows

- Object: triangle and line segment
- Concept: perpendicular
- Theorem: similarity, trigonometry and Inequality
- Number of questions: 1

The input given in the first example covers the concept of *perpendicular lines*. The generated questions are required to make use of theorems involving *similarity of triangles, trigonometry and inequality*. Figure 16 shows the step by step execution of the algorithm for the given input. Figure 16a generates a configuration involving *triangle and perpendicular* geometry objects using the GF component. This figure is generated by non-deterministically selecting a way of generating perpendicular lines in a triangle from our predefined database.

Figure 16b is generated using GD component with the help of KT in the figure 16a. There are various ways of making two triangles similar. Firstly, this component performs pattern matching to find the suitable ways of making two triangles similar in the existing figure. Later, non-deterministically, it picks one way of making triangles similar and generates data accordingly. Now GD component assigns data using the trigonometry rules stored in the database. Currently, the database is limited to the basic trigonometry rules related to the sides and angles of a triangle. For example, "$\sin 2\theta = 2\sin\theta\cos\theta$" trigonometry rule does not includes any sides and angles in a triangle. Hence, it is not included in our database.

Figure 16c shows the new facts and the their reasoning generated using the GFS component. The new fact is the new question and the reasoning is the solution of the generated question.

5.2 Second Example

The input given in this question generation task is as follows

- Object: triangle and line segment
- Concept: parallel line
- Theorem: similarity and ratio of length of sides
- Number of questions: 1

The input given in the second example covers the concept of *parallel line*. The generated questions are required to make use of theorems involving *similarity and ratio of length of sides* of a triangle. Figure 17 shows the step by step execution of the algorithm for

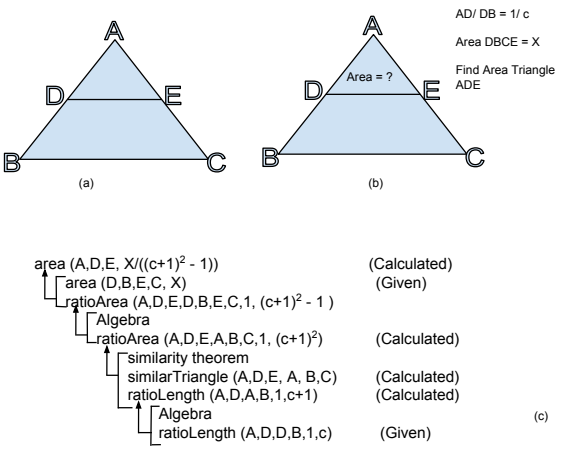


Figure 17: (a) The figure is generated by using *triangle* and *parallel line* as the geometry objects using the GF component. (b) The data involving ratio of length of sides is generated for (a) using the GD component. (c) The new fact and its derivation using the GFS component is generated from the figure shown in (b).

the given input. Figure 17a generates a configuration involving *triangle and parallel line* geometry objects using the GF component. Before generating similarity of two triangles, it finds similar triangles in the existing configuration. It can be seen that the existing configurations has similar triangles ADE and ABC. Hence, there is no need to generate similar triangles. Application of similarity theorem results in the generation of ratios of sides of length of two triangles. Figure 17b is generated using GD component with the help of KT knowledge database. Figure 17c shows the new facts and the their reasoning generated using the GFS component. The new fact is the new question and the reasoning is the solution of the generated question.

5.3 Third Example

The input given in this question generation task is as follows

- Object: triangle and line segment
- Concept: perpendicular and angle bisector
- Theorem: Sum of angle in a triangle
- Number of questions: 1

The input given in the third example covers the concept of *angle-bisector*. The generated questions are required to make use of theorems about sum of angles in a triangle. Figure 18 shows the step by step execution of the algorithm for the given input. Figure 18a generates a configuration involving *triangle and perpendicular* geometry objects using the GF

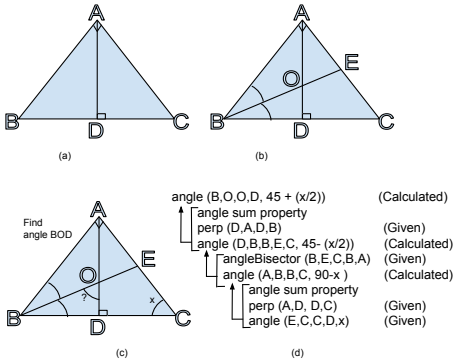


Figure 18: (a) The figure is generated by using *triangle* and *perpendicular* as the geometry objects using the GF component. (b) *Angle-bisector* is added to (a) using the GF component.(c) The data involving angles in a triangle is generated for (b) using the GD component. (d) The new fact and its derivation using the GFS component is generated from the figure shown in (c).

component. Figure 18b generates a new configuration from (a) involving *angle-bisector* geometry object. This figure is generated by non-deterministically selecting a way of generating angle-bisector in a triangle. The decision of generating perpendicular before angle-bisector is taken randomly. Figure 18c is generated using GD component with the help of KT knowledge database. There are many triangles in the existing configuration. Hence, there are various ways of assigning data to satisfy angle-sum property in these triangles. GD component picks one triangle and generates the data accordingly. Figure 18d shows the new facts and the their reasoning generated using the GFS component. The new fact is the new question and the reasoning is the solution of the generated question.

6 IMPLEMENTATION

Each component of our tool is implemented independently, using state-of-the-art libraries and systems. Various programming languages are used in the system. C++ is used for performing calculations and Python is used for implementation of the algorithms used in GF and GD components. The algorithm in GFS component is implemented using Constraint Handling Rules (CHR) (Frühwirth and Raiser, 2011). CHR are used for generating new facts from the axioms and the given facts. In our implementation, we use the CHR library provided by K.U.Leuven, on top of SWI-Prolog (Schrijvers and Demoen, 2004). The theorems used in GFS component are manually converted in the format used by CHR library. For implementing knowledge representation KF and KT, the graph database Neo4j (Vicknair and Macias, 2010) is

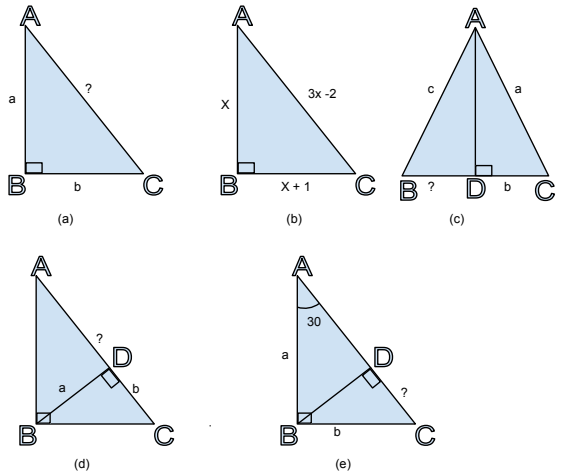


Figure 19: Generated questions based on "triangle", "perpendicular" and "Pythagorean Theorem" as input. Questions in figure(a-c)can be solved using Pythagorean Theorem only. However, questions of figure (d) requires similarity and figure (e) requires trigonometry for finding the unknown value.

used. KF knowledge graph is used and modified by all the components and finally represents the question. Our knowledge databases such KT and predefined ways of intersection of geometric objects are manually generated and stored before the questions generation.

6.1 Experimental Results

The system can generate geometry questions using the framework described in Section 4. Currently, our knowledge database of objects contains line segments and triangles. In addition, we have a predefined set of more than 100 theorems. The generated questions cover five categories, e.g. similarity. Figure 19 shows various questions generated by our system on selecting "*triangle*" as object, "*perpendicular*" as concept and "*Pythagorean Theorem*" as a theorem to be covered. The generated questions are tested using the existing geometry solver tool JGEX (Gao and Lin, 2004). For testing in JGEX, the figure configuration is drawn manually by the user and the system is asked to prove/find a certain relationship. The tool is able to prove/solve all the questions generated by our system. Comparing the solutions generated by JGEX with the solutions generated by our system, we found interesting differences that may stem from different representations of geometric knowledge and reasoning techniques and that deserve further investigation.

7 CONCLUSION

In this paper, we provide a framework for the automatic generation and solving of questions for high school mathematics, specifically in the geometry domain. Our system is able to quickly generate large numbers of questions on specific topics. Such a system will help teachers reduce the time and effort spent on the tedious and error-prone task of generating questions. Our work aims to develop an automated geometry question generation system that uses a deductive approach for finding the relations between mathematical concepts and for generating and proving these conjectures about concepts.

Future work can be carried in various directions. An experiment needs to be performed in which the generated problems would be placed in front of teachers and let them guess which problem is from a textbook and which one was generated from our algorithm. Other major work would be generating questions according to the required difficulty level. Another improvement would be the addition of knowledge by the user. Lastly, a GUI should be developed in order to make it usable for teachers, students and testers.

REFERENCES

Chou, S. C. and Gao, X. S. (2001). *Handbook-of-Automated-Reasoning*. Elsevier and MIT Press.

Chou, S. C., Gao, X. S., and Zhang, J. Z. (1994). Machine proofs in geometry: Automated production of readable proofs for geometry theorems.

Chou, S. C., Gao, X. S., and Zhang, J. Z. (2000). A deductive database approach to automated geometry theorem proving and discovering. *J. Autom. Reason.*, 25(3):219–246.

Chou, S. C., Gao, X. S., and Zhang, J. Z. (2011). Area method and automated reasoning in affine geometries.

Coelho, H. and Pereira, L. (1986). Automated reasoning in geometry theorem proving with prolog. *Journal of Automated Reasoning*, 2(4):329–390.

Francis, G. (2002). Axiomatic systems for geometry.

Frühwirth, T. and Raiser, F., editors (March 2011). *Constraint Handling Rules: Compilation, Execution, and Analysis*.

Gao, X.-S. and Lin, Q. (2004). Mmp/geometer asoftware package for automated geometric reasoning. In Winkler, F., editor, *Automated Deduction in Geometry*, volume 2930 of *Lecture Notes in Computer Science*, pages 44–66. Springer Berlin Heidelberg.

Gao, X. S., Zhu, C. C., and Huang, Y. (1998). Building dynamic mathematical models with geometry expert, i. geometric transformations, functions and plane curves. In *Proceedings of the Third Asian Technology Conference in Mathematics*.

Greeno, G., J., Magone, and Maria E. Chaiklin, S. (1979). Theory of constructions and set in problem solving. *Memory and Cognition*, 7:445–461.

Jiang, J. and Zhang, J. (2012). A review and prospect of readable machine proofs for geometry theorems. *Journal of Systems Science and Complexity*, 25:802–820.

McCharen, J. D., Overbeek, R. A., and Wos, L. A. (1976). Problems and experiments for and with automated theorem-proving programs. *IEEE Trans. Comput.*, 25(8):773–782.

Meikle, L. and Fleuriot, J. (2003). , formalizing hilberts grundlagen in isabelle/isar. In *Theorem Proving in Higher Order Logics*.

Narboux, J. (2010). The area method: a recapitulation. *Journal of Automated Reasoning*, 48:489–532.

National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics.

Nevins, A. (1975). Plane geometry theorem proving using forward chaining. *Artificial Intelligence*, 6(1):1–23.

Quaife, A. (1989). Automated development of tarskis geometry. *Journal of Automated Reasoning*, 5:97–118.

Schrijvers, T. and Demoen, B. (2004). The k.u.leuven chr system: Implementation and application.

Vicknair, C. and Macias, M. (2010). *A Comparison of a Graph Database and a Relational Database*.

Wilson, S. and Fleuriot, J. D. (2005). Combining dynamic geometry, automated geometry theorem proving and diagrammatic proofs. In *European joint conference on Theory And Practice of Software - ETAPS*.

Zhang, J. Z., Gao, X. S., and Chou, S. C. (1996). The geometric information search system by forward reasoning. *Chinese Journal of Computers*, 19(10):721727.

Finding Regularities in Courses Evaluation with K-means Clustering

R. Campagni, D. Merlini and M. C. Verri
Dipartimento di Statistica, Informatica, Applicazioni, Università di Firenze
Viale Morgagni 65, 50134, Firenze, Italia
{*renza.campagni, donatella.merlini, mariacecilia.verri*}*@unifi.it*

Keywords: Educational Data Mining, K-means Clustering, Courses Evaluation, Assessment.

Abstract: This paper presents an analysis about the courses evaluation made by university students together with their results in the corresponding exams. The analysis concerns students and courses of a Computer Science program of an Italian University from 2001/2002 to 2007/2008 academic years. Before the end of each course, students evaluate different aspects of the course, such as the organization and the teaching. Evaluation data and the results obtained by students in terms of grades and delays with which they take their exams can be collected and reorganized in an appropriate way. Then we can use clustering techniques to analyze these data thus show possible correlation between the evaluation of a course and the corresponding average results as well as regularities among groups of courses over the years. The results of this type of analysis can possibly suggest improvements in the teaching organization.

1 INTRODUCTION

The evaluation of university education is an important process whose results can be used in the programming and management of the educational activities by monitoring resources (financial, human, structural and others), services (orientation for students and administrative offices), students careers, courses and occupancy rate. In order to evaluate all these aspects, it is important to analyse the opinion of the *users* of university education, i.e. the students.

The evaluation of the learning process falls in the context of the Educational Data Mining (EDM), an emerging and interesting research area that aims to identify previously unknown regularities in educational databases, to understand and improve student performance and the assessment of their learning process. As described in (Romero and Ventura, 2010), EDM uses statistical, machine learning and data mining algorithms on different types of data related to the field of education. It is concerned with developing methods for exploring these data to better understand the students and the frameworks in which they learn thus possibly enhancing some aspects of the quality of education. Data mining techniques have also been applied in computer-based and web-based educational systems (see, e.g., (Romero et al., 2010; Romero et al., 2008)). In this paper, we use a data mining approach based on `K-means` clustering to link the evaluation of courses taken by students with their results, in terms of average grade and delay in the corresponding exams. We also analyse the evaluation of courses over the years in order to identify similar behaviors or particular trends among courses, by using an approach similar to time series clustering (see, e.g., (Liao, 2005)).

This study deepens the analysis presented in (Campagni et al., 2013) and is analogous to that used in (Campagni et al., 2012a; Campagni et al., 2012b; Campagni et al., 2012c). The analysis refers to a real case study concerning an Italian University but it could be applied to different scenarios, except for a possible reorganization of the involved data. The data set is not very large but allows us to illustrate a quite general methodology on a real case study. Our approach uses standard data mining techniques, but we think very interesting the concrete possibility of applying these techniques to find and analyse patterns in the context of university courses evaluation, even in large universities.

2 DATA FOR ANALYSIS

In this section, we describe how courses are evaluated by students at the University of Florence, in Italy, with the aim of providing a methodology to search for regularities in data concerning courses evaluation. Therefore, the steps we present can be ap-

plied also in other academic contexts. In particular, we refer to a Computer Science degree of the Science School, under the Italian Ministerial Decree n. 509/1999. This academic degree was structured over three years and every academic year was organized in two semesters; there were several courses in each of these six semesters and at the end of a semester students could take their examinations. Exams could be taken in different sessions during the same year, after the end of the corresponding courses, or later.

Table 1 illustrates an example of students data after a preprocessing phase which allow us to integrate original attributes, such as the grade and the date of the exam, with both the semester in which the course was given, **Semester1**, and the semester in which the exam was taken, **Semester2**. Finally, we can compute the value **Delay** as the difference between the semester of the course and the semester in which the student took the exam. We highlight that the values of attributes **Semester1** and **Semester2** are not usually stored in the databases of the universities, therefore this preprocessing phase may be onerous.

At the University of Florence, starting from the academic year 2001/2002, a database stores information about evaluation of the courses quality of various degree programs, among which we find the degree under consideration. The results of this process are available at the address (SISValDidat), under permission of the involved teacher, and show for each course several pieces of information, such as the name of the teacher who took the course and the average rating given by students on various topics. Before the end of each course (at about 2/3 of the course), students compile, anonymously, a module to express their opinion on the course just taken. This form is divided into the following five paragraphs:

- paragraph 1, concerns the organization of the degree program;
- paragraph 2, concerns the organization of the course;
- paragraph 3, concerns the teacher;
- paragraph 4, concerns classrooms and equipment;
- paragraph 5, concerns the general satisfaction about the course.

Each paragraph is composed by some questions; students can choose among four levels of answers, two negative and two positive levels (disagree, slightly disagree, slightly agree, agree). For details the interested reader can see the sample of the module in (SISValDidat).

For each course of an academic year and for each paragraph, we can compute the percentage of positive answers, that is, of type *slightly agree* and *agree* by grouping together all questions belonging to the same paragraph and their average percentage value.

To relate data of students careers with courses evaluation, for each course we can compute the average grade and the average delay attained by students who took the exam in the same year. An example of this data organization is illustrated in the first four columns of Table 2. As already observed, the evaluation of courses is anonymous and is done only by students who really take the course, therefore, in this kind of organization, it may happen to consider information concerning exams of students who may not be the same students who evaluated the courses. As a consequence, we can only compare the results of courses evaluation in a specific year with the aggregate results of students who took the corresponding exams in the same period. However, this data organization does not change a lot if it was possible to identify the students involved in the courses evaluation in order to connect properly the results of the evaluation with those of exams. Obviously, in this case we should ensure the privacy of results, for example by using a differential privacy approach (see, e.g., (Dwork, 2008)).

After a preprocessing phase, we can organize students and evaluation data into two different ways by taking into account the following fields:

- **Exam**, the code which identifies an exam;
- **Year**, the year of the evaluation;
- **AvgGrade**, the average grade of the exam;
- **AvgDelay**, the average delay, in semesters, of students exams;
- **Par*k*(t)**, the percentage of positive evaluations of paragraph *k* at time *t*.

In particular, Table 2 illustrates a sample of the dataset which can be used to compare examination results and courses evaluation while Table 3 represents a sample of data that can be used to analyze the evolution over the years of courses evaluation. As we will illustrate in Section 3, data organized as in Table 2 will be clustered with `K-means` algorithm by using the Euclidean distance to separate the *multidimensional points* representing some characteristic of a course in a specific year; data organized as in Table 3 will be represented in the plane as *trajectories* corresponding to the evaluation of courses over the years and will be clustered with the Manhattan distance. Both these approaches can be used to find regularities in courses evaluations and can highlight criticalities or suggest improvements in the teaching organization.

Table 1: A sample of students data: grades in thirtieths.

Student	Exam	Date	Grade	Semester1	Semester2	Delay
100	10	2001-01-14	24	1	1	0
100	20	2002-12-20	27	2	3	1
200	20	2002-06-04	21	2	2	0
300	10	2001-01-29	26	1	3	2
400	10	2002-02-15	26	1	2	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Data organization for comparing examination results and courses evaluation.

Exam	Year	AvgGrade	AvgDelay	Par*1*	...	Par*5*
10	2001	25	1	51	...	60
10	2002	26	1	50	...	61
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	2007	25	1	81	...	67
20	2001	24	0.5	56	...	77
20	2002	26	1	62	...	59
⋮	⋮	⋮	⋮	⋮	⋮	⋮

3 K-MEANS CLUSTERING WITH EUCLIDEAN AND MANHATTAN DISTANCES

Among the different data mining techniques, clustering is one of the most widely used methods. The goal of cluster analysis is to group together objects that are similar or related and, at the same time, are different or unrelated to the objects in other clusters. The greater the similarity (or homogeneity) is within a group and the greater the differences between groups are the more distinct the clusters are. `K-means` is a very simple and well-known algorithm based on a partitional approach; it was introduced in (MacQueen, 1967) and a detailed description can be found in (Tan et al., 2006). In this algorithm, each cluster is associated with a centroid and each point is assigned to the cluster with the closest centroid by using a particular distance function. The centroids are iteratively computed until a fixed point is found. The number K of clusters must be specified. In particular, in this paper we use both the *Euclidean* and *Manhattan* distance; in the first case, the centroid of a cluster is computed as the mean of the points in the cluster while in the second case the appropriate centroid is the median of the points (see, e.g., (Tan et al., 2006)).

The evaluation of the clustering model resulting from the application of a cluster algorithm is not a well developed or commonly used part of cluster analysis; nonetheless, cluster evaluation, or cluster validation, is important to measure the goodness of the resulting clusters, for example to compare clustering algorithms or to compare two sets of clusters. In our analysis we measured cluster validity with correlation, by using the concept of proximity matrix and incidence matrix. Specifically, after obtaining the clusters by applying `K-means` to a dataset, we computed the proximity matrix $P = (P_{i,j})$ having one row and one column for each element of the dataset. In particular, each element $P_{i,j}$ represents the Euclidean, or Manhattan, distance between elements i and j in the dataset. Then, we computed the incidence matrix $I = (I_{i,j})$, where each element $I_{i,j}$ is 1 or 0 if the elements i and j belong to the same cluster or not. We finally computed the Pearson's correlation, as defined in (Tan et al., 2006, page 77), between the linear representation by rows of matrices P and I. Correlation is always in the range -1 to 1, where a correlation of 1 (-1) means a perfect positive (negative) linear relationship.

As a first example, Table 4 illustrates the final grade and the graduation time, expressed in years, of a sample of graduated students. By applying the `K-means` algorithm to this dataset, with $K = 2$, **FinalGrade** and **Time** as clustering attributes and by using the Euclidean distance, we obtain the following two clusters, in terms of the student identifiers: $C_1 = \{100, 400, 600, 700\}$ and $C_2 = \{200, 300, 500\}$; the centroids of the clusters have coordinates $\mathcal{C}_1 = (107, 3.5)$ and $\mathcal{C}_2 = (96, 5.33)$, respectively. Tables 5 and 6 show the proximity matrix and the incidence matrix corresponding to clusters C_1 and C_2 of the data set illustrated in Table 4. The Pearson's correlation between the linear representation of these two matrices is -0.59, a medium value of correlation.

Table 3: Data organization for analyzing the trend over the years of courses evaluation.

Exam	**Par*1*(2001)**	...	**Par*1*(2007)**	...	**Par*5*(2001)**	...	**Par*5*(2007)**
10	51	...	81	...	60	...	67
20	56	...	84	...	77	...	84
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 4: A sample data set about students.

Student	**FinalGrade**	**Time**
100	110	3
200	95	5
300	100	5
400	103	4
500	98	6
600	106	4
700	109	3

Table 5: The proximity matrix for data of Table 4.

P	**100**	**200**	**300**	**400**	**500**	**600**	**700**
100	0						
200	20.12	0					
300	10.25	10	0				
400	7.07	13.08	3.32	0			
500	12.41	8.06	2.24	5.48	0		
600	4.12	16.06	6.16	3	8.31	0	
700	1	19.13	9.27	6.08	11.45	3.16	0

As another example, Table 7 shows a sample of data concerning courses evaluation: in particular, each row contains the exam identifier and the percentage of positive evaluation of a generic paragraph at time t_i, for $i = 1,\cdots,4$. We can apply the `K-means` algorithm to the dataset in Table 7, with $K = 2$, **Par(t_i)**, for $i = 1,\cdots,4$, as clustering attributes and by using the Manhattan distance. This means to represent each element of the data set as a broken line connecting the points (t_i,**Par(t_i)**), for $i = 1,\cdots,4$, in the cartesian plane. The Manhattan distance between two broken lines thus corresponds to the sum of the vertical distances between the ordinates. By using the `K-means` algorithm, we obtain the following two clusters in terms of course identifiers: $C_1 = \{200,500\}$ and $C_2 = \{100,300,400\}$; the centroids of the clusters are represented by the sequences $\mathcal{C}_1 = [(1,72),(2,68),(3,67),(4,73)]$ and $\mathcal{C}_2 = [(1,82.5),(2,91),(3,87.5),(4,91.5)]$, respectively. Figure 1 illustrates the clustering result by evidencing the centroids $\mathcal{C}_1$ and $\mathcal{C}_2$.

Table 6: The incidence matrix for clustering of data of Table 4.

I	**100**	**200**	**300**	**400**	**500**	**600**	**700**
100	1						
200	0	1					
300	0	1	1				
400	1	0	0	1			
500	0	1	1	0	1		
600	1	0	0	1	0	1	
700	1	0	0	1	0	1	1

Table 7: A sample data set about courses evaluation.

Exam	**Par(t_1)**	**Par(t_2)**	**Par(t_3)**	**Par(t_4)**
100	55	65	67	60
200	85	87	85	92
300	72	68	65	77
400	77	80	70	73
500	80	95	90	91

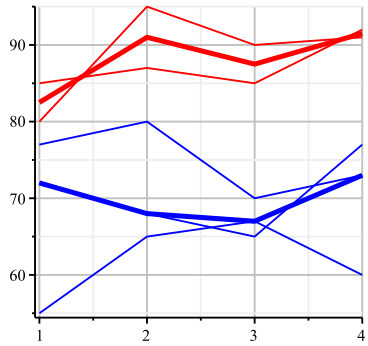


Figure 1: `K-means` results with data of Table 7 with $K = 2$ and Manhattan distance, centroids in evidence.

Also in this case we can compute the Pearson's correlation by using the proximity and the incidence matrices computed by using the Manhattan distance.

3.1 The Case Study

As already observed, the real datasets we analysed concern courses and exams during the academic years from 2001/2002 to 2007/2008 at the Computer Science program of the University of Florence, in Italy. In particular, the first data set is organized as illustrated in Table 2 and refers to the evaluation of 40 courses in seven different years. We explicitly observe that we did not consider in our analysis those courses evaluated by a small number of students. For clustering, we used the `K-means` implementation of

Weka (Witten et al., 2011), an open source software for data mining analysis. The aim was to find if there is a relation between the valuation of a course and the results obtained by students in the corresponding exam. We performed several tests with different values of the parameter K and we selected different groups of attributes. We point out that the attributes selection is an important step and should be done according to the preference of an expert of the domain, for example the coordinator of the degree program. For each choice of attributes, we applied the K-means algorithm with the Euclidean distance to identify the clusters; then, we computed the Pearson's correlation by using the proximity and incidence matrices. The tests we performed pointed out that the exams having good results, in terms of average grade and delay, correspond to courses having also a good evaluation from students.

In particular, we used **AvgGrade**, **AvgDelay**, **Par*1***, **Par*2***, **Par*3***, **Par*4*** and **Par*5*** as clustering attributes and $K = 2$, obtaining the clusters illustrated in Figures 2, 3, 4 and 5; each figure represents the projection of the clusters along two dimensions corresponding to the following pairs of attributes **AvgDelay** and **Par3**, **AvgGrade** and **Par3**, **AvgDelay** and **Par4** and, finally, **AvgGrade** and **Par4**. The centroids of the resulting clusters are shown in Table 8, which also contains the average values relative to the full data set.

Table 8: The centroids of clusters in Figures 2, 3, 4, 5.

Attribute	Full Data	Cluster0	Cluster1
AvgGrade	25.31	25.85	24.58
AvgDelay	2.61	1.8	3.68
Par*1*	70.86	77.74	61.67
Par*2*	72.23	82.19	58.94
Par*3*	84.51	90.25	76.86
Par*4*	72.03	74.67	68.5
Par*5*	76.02	80.83	69.61

The cluster number 0, which correspond to 88 blue stars in the figures, contains the courses which students took with *small* delay and that they evaluated positively. On the other hand, cluster number 1, corresponding to 66 red stars, contains those courses which students took with a large delay and that they evaluated less positively. We observe that the centroids of the two clusters are very close relative to the attribute **Par4** which concerns classrooms and equipment. This is also evidenced from Figures 4 and 5, where the blue and red stars are less separated than those in Figures 2 and 3. The Pearson's correlation corresponding to these clusters is equal to -0.35. We obtained an improvement by excluding the attribute

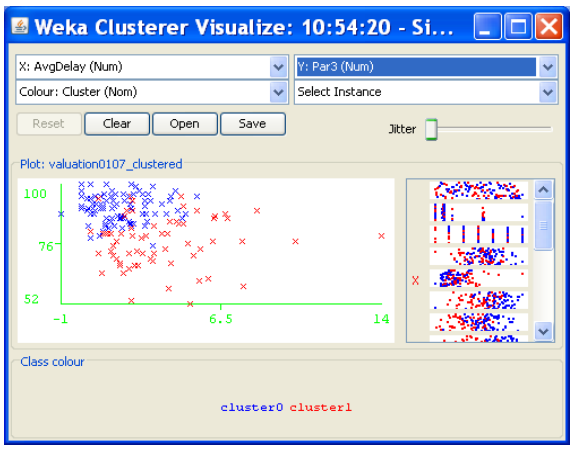


Figure 2: Clusters of Table 8 with **AvgDelay** and **Par3** in evidence.

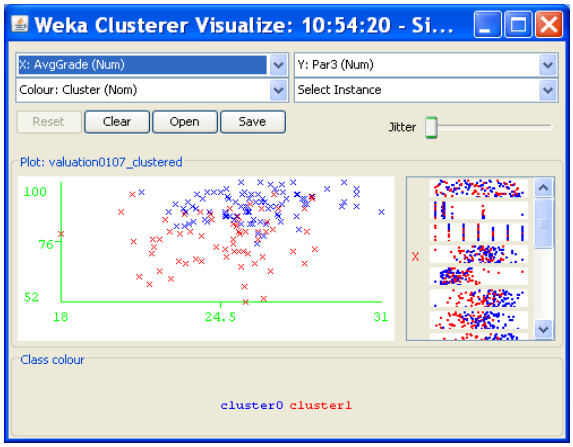


Figure 3: Clusters of Table 8 with **AvgGrade** and **Par3** in evidence.

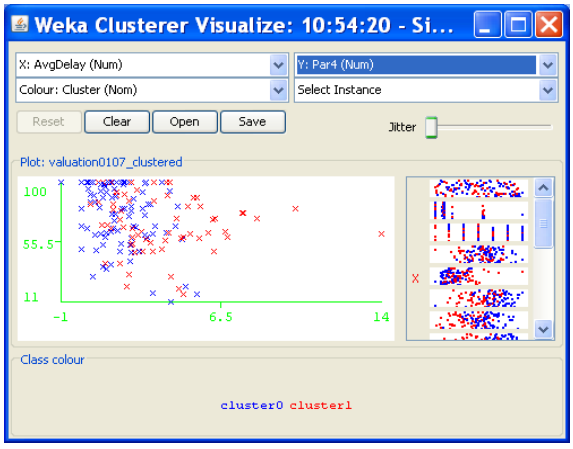


Figure 4: Clusters of Table 8 with **AvgDelay** and **Par4** in evidence.

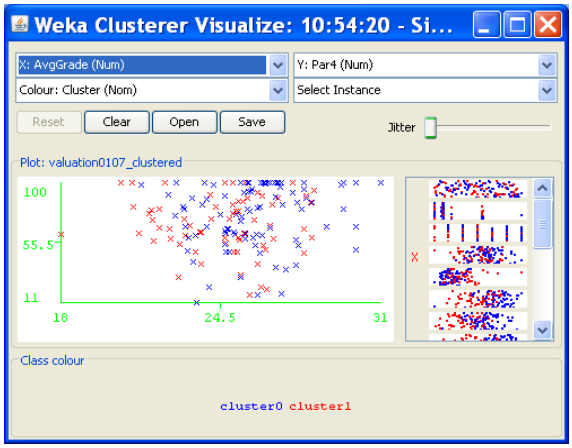


Figure 5: Clusters of Table 8 with **AvgGrade** and **Par4** in evidence.

Par4 from clustering, in fact in this case we find a correlation equal to -0.51. In general, our tests evidenced that the paragraphs evaluations which are more correlated with students results regard attributes **Par*2*** and **Par*3***, that is, those concerning the course organization and the teacher. We point out that the value $K = 2$ gave the best results in terms of correlation.

Among the courses considered in the previous data set, we selected those evaluated all seven years, for a total of sixteen courses, some in Mathematics and others in Computer Science. This time we are interested in analysing data organized as in Table 3, by considering the evaluation of a particular paragraph over the years. The aim was to find if there are similar behaviors among courses, that is, if we can classify courses according to their evaluations. We performed several tests, by choosing a paragraph at a time. For each choice of attributes, we applied the `K-means` algorithm with the Manhattan distance to identify the clusters; also in this case we computed the Pearson's correlation by using the proximity and incidence matrices.

Figure 6 illustrates the result of `K-means` with $K = 2$, Manahattan distance and **Par*2*(2001)**, **Par*2*(2002)**, $\cdots$, **Par*2*(2007)** as clustering attributes. The points defining the centroid trajectories of the resulting clusters are shown in Table 9, which also contains the median values relative to the full data set. The Pearson's correlation corresponding to these clusters is equal to -0.64.

The figure puts well in evidence that the courses are divided into two clusters with well distinct centroids. The red cluster contains courses that have been evaluated better over the years while the blue cluster corresponds to courses that students rated worse.

Table 9: The points defining the centroid trajectories of clusters in Figure 6.

Attribute	Full Data	Cluster0	Cluster1
Par*2*(2001)	73.5	85	57
Par*2*(2002)	77	87	56
Par*2*(2003)	73.5	84	52
Par*2*(2004)	72.5	79	58
Par*2*(2005)	74.5	79	65
Par*2*(2006)	78.5	84	71
Par*2*(2007)	75.5	83	69

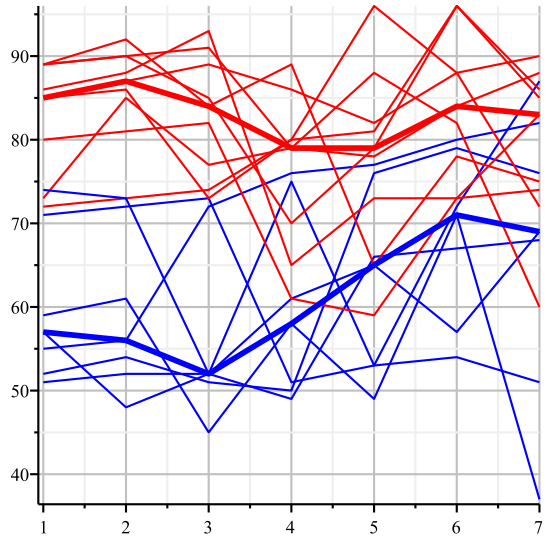


Figure 6: Clusters of Table 9 with centroids in evidence: each line represents the percentage of positive evaluations about the organization of a course (paragraph 2) over the years 2001-2007.

What is interesting, though not surprising, is that all courses in the red cluster are Computer Science courses while the blue cluster contains many Mathematics courses. We highlight that the centroids show rather clearly the behavior of the assessment over the years. In particular, the evaluation of the courses in the blue cluster has improved over the years while that of courses in the red cluster has remained more stable.

Also in this case the best results in terms of correlation were found with $K = 2$; however, with $K = 4$ we found the courses rated worse distributed into two clusters, one of which contains only the Mathematics courses. The corresponding centroid illustrates a gradual improvement of the assessment for this type of courses during the years under examination.

4 CONCLUSION AND FUTURE WORK

The results of the previous sections show, in a formal way with data mining techniques, that there is a relationship between the evaluation of the courses from students and the results they obtained in the corresponding examinations. In particular, the analysis performed on data related to the Computer Science degree program under examination illustrates that the courses which received a positive evaluation correspond to exams in which students obtained a good average mark and that they took with a small delay. Conversely, the worst evaluations were given to those courses which do not match good achievements by students.

The analysis based on clustering with Manhattan distance allows us to classify courses according to the assessment received by students and can highlight some regularities that emerge over the years or points out some trend reversals due to changes of teachers. In the Computer Science degree program just considered, for example, we observe the trend to give not so good evaluation to Mathematics courses. Results of this type point out a critical issue in the involved courses and can be used to implement improvement strategies.

We wish to emphasize that our analysis refers to the courses evaluation that students make before taking the exams and knowing their grades. In fact, as already observed, the evaluation module is given to students before the end of the course. Surely, there is the risk that their judgment is influenced by the inherent difficulty of the course or by the comments made by students of the previous years. To this purpose, it is important that during the module compilation the teacher explains that a serious assessment of the course can increase the quality level of the involved services. Students represent the end-users as well as the principal actors of the formative services offered by the University and the measure of their perceived quality is essential for planning changes. However, the results of courses evaluation should always be considered in a critical way and should not have the goal of simplifying the contents to get best ratings.

In general, many other factors should be considered for evaluating courses and student success, as addressed in (Romero and Ventura, 2010). The approach used in this work could be refined and deepened if it was possible to identify the students involved in the courses evaluation in order to connect properly the results of the evaluation with those of exams. Moreover, it would be interesting to connect the assessment of students with other information such as the gender of students and teachers or the kind of high school attended by students. Starting from the academic year 2011/2012, the University of Florence began to manage on line the evaluation module described in Section 2. Therefore, in a next future, it might be possible to proceed in this direction, taking into account appropriate strategies to maintain privacy.

An interesting additional source of information could be given by social media sites, such as *Facebook* or *Twitter*, used by students to post comments about courses and teachers. It would be useful to link this information with the results of students and their official evaluations about teachings, in order to take into account more feedbacks. In such a context, it might be interesting to use text mining techniques to classify the student comments and enrich the database for an analysis similar to that illustrated in this work.

REFERENCES

Progetto SISValDidat. https://valmon.disia.unifi.it/sisvaldidat/unifi/index.php.

Campagni, R., Merlini, D., and Sprugnoli, R. (2012a). Analyzing paths in a student database. In *The 5th International Conference on Educational Data Mining, Chania, Greece*, pages 208–209.

Campagni, R., Merlini, D., and Sprugnoli, R. (2012b). Data mining for a student database. In *ICTCS 2012, 13th Italian Conference on Theoretical Computer Science, Varese, Italy*.

Campagni, R., Merlini, D., and Sprugnoli, R. (2012c). Sequential patterns analysis in a student database. In *ECML-PKDD Workshop: Mining and exploiting interpretable local patterns (I-Pat 2012), Bristol*.

Campagni, R., Merlini, D., Sprugnoli, R., and Verri, M. C. (2013). Comparing examination results and courses evaluation: a data mining approach. In *Didamatica 2013, Pisa, Area della Ricerca CNR, AICA*, pages 893–902.

Dwork, C. (2008). Differential privacy: a survey of results. In *Theory and Applications of Models of Computation, 5th International Conference, TAMC 2008*, pages 1–19.

Liao, T. W. (2005). Clustering of time series data: a survey. *Pattern Recognition*, 38(11):1857–1874.

MacQueen, J. (1967). Some methods for classifications and analysis of multivariate observations. In *Proc. of the 5th Berkeley Symp. on Mathematical Statistics and Probability. University of California Press.*, pages 281–297.

Romero, C., Romero, J. R., Luna, J. M., and Ventura, S. (2010). Mining rare association rules from e-learning data. In *The 3rd International Conference on Educational Data Mining*, pages 171–180.

Romero, C. and Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on systems, man and cybernetics*, 40(6):601–618.

Romero, C., Ventura, S., and García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1):368–384.

Tan, P. N., Steinbach, M., and Kumar, V. (2006). *Introduction to Data Mining*. Addison-Wesley.

Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*. Morgan Kaufmann.

Cross-analysis of Transversal Competences in Project Management

Ana González-Marcos[1], Fernando Alba-Elías[1], Joaquín Ordieres-Meré[2] and Fermín Navaridas-Nalda[3]
[1]*Department of Mechanical Engineering, Universidad de La Rioja, c/ Luis de Ulloa 20, 26004 Logroño, La Rioja, Spain*
[2]*PMQ Research Group, ETSII, Universidad Politécnica de Madrid, José Gutiérrrez Abascal 2, 28006 Madrid, Spain*
[3]*Department of Education Sciences, Universidad de La Rioja, c/ Luis de Ulloa 8, 26004 Logroño, La Rioja, Spain*
{ana.gonzalez, fernando.alba}@unirioja.es, j.ordieres@upm.es, fermin.navaridas@unirioja.es

Keywords: Project Management, Transversal Competences, Competence Assessment.

Abstract: This paper presents a framework for project management competence assessment based on participant's performance and contribution in a simulated environment. The proposed framework considers competence assessment through evidences and the participation of different roles. The system enforces the assessment of individual competences by means of a set of performance indicators. A specific case study is presented and the relationship between exhibited transversal competences and project quality is analysed.

1 INTRODUCTION

The European space for higher education is immersed in a substantial transformation process of the organizational, pedagogical and methodological aspects of knowledge transmission (Council of the European Union, 2004). In this changing context, the new educational model focuses on learning and competences' development.

In a constantly changing society, the demands faced by an individual vary from one situation to another and from time to time. Therefore, in addition to possessing the specific basic skills for accomplishing a certain task, more flexible, generic and transferable competences are needed to provide the individual with a combination of skills, knowledge and attitudes that are appropriate to particular situations (European Commision, 2004).

In this new higher education conception, the instructor's role shifts from transmitting knowledge to students to facilitate and guide their learning process (Beltrán, 1999; Navaridas, 2004). Thus, the teaching process must be organized in a more learner-centered approach than classical lectures offer.

This work presents a teaching framework that aims to stimulate the learning of both technical and human skills in project management. More specifically, this study concentrates on the relationship observed between four transversal competences –teamwork, leadership, motivation and results orientation- and the final project success.

The structure of the remaining part of the paper is as follows. Section 2 presents a brief review on related works. Section 3 provides an overview of the learning experience and section 4 is dedicated to present and discuss the relationships observed between the four analysed competences and the final project quality. Finally, the last section discusses some general conclusions and presents future work.

2 BACKGROUND

Competences in project management are correlated to performance on the job and can be confronted against well-accepted standards and improved via training and other development activities. The underlying assumption is that a competency can be broken down into dimensions of competence, as Project Management Competency Development (PMCD) Framework does. In this framework, considered dimensions are Knowledge, Personal and Performance (PMI, 2002). Furthermore, the International Project Management Association (IPMA) created the International Competence Baseline (ICB) consisting of 46 elements for knowledge and experience as well as personal attitudes and abilities for general impression (IPMA, 2006). In addition, the Association for Project Management (APM) developed the APM Competence Framework providing the technical, behavioural and contextual competence elements needed for effective project management (APM,

2008).

Even though human skills –communication, teamwork, organizational effectiveness, leadership, flexibility, creativity, etc.– are acknowledged as important for project management, the education offered in industrial engineering degrees concentrates mainly on the control aspects of projects, i.e., the technical skills. It is recently that authors have started to discuss how to teach this discipline in higher education. Thus, Pant and Baroudi (2008) argue the necessity of a more balanced approach between technical and human concepts to enhance project management education. Clark (2008) discusses the skills required for an effective project manager, as well as the analysis of four approaches at the M.Sc. level to develop these skills. Barron (2005) discusses the difficulty of learning effective project management skills and suggests that there is a way to teach project management through properly designed assessment. In the same way, Sense (2007) emphasizes that project learning and the learning of behaviours that will lead to success are most appropriately pursued through the creation of a suitable environment.

The teaching and learning of project management has grown in interest and popularity (Berggren and Söderlund, 2008; Ojiako et al., 2011; Reif and Mitri, 2005) and there are some practical approaches to the teaching of project management. For instance, Abernethy et al. (2007) describe a specific experimental approach for information technology students. Authors argue that project activities must mirror the real world for information technology students to learn what needs to be done in industry projects. More recently, Crespo et al. (2011) advocated a combination of theoretical content, individual applied tasks, use of software systems and a strategy of learning by doing in teaching project management. They formally introduce negotiating and virtual team management aspects to different teams from different universities in different locations.

3 METHODOLOGY AND IMPLEMENTATION

The Project and Portfolio Management Learning (P2ML) framework presented in this work is based on the *experiential learning* theory, i.e., learning through action. The following is a list of the main characteristics of the proposed model:

- Students are involved in real-world engineering projects, which provides authenticity and require students to use academic and technical knowledge.
- Students are forced to adopt a more active role since they are the ones who must develop a project within given time and specifications.
- Acquisition of teamwork abilities and human skills, such as leadership, communication or negotiation, are promoted.
- Professional skills that should be deployed by a project manager are implemented in scale.

In terms of the selected method for project management, it was adopted PRINCE2™ (Projects IN a Controlled Environment) (Office of Government Commerce, 2009) as far as it is simple, product oriented and easier to understand for students without any previous experience in neither projects nor project management.

The use of PRINCE2™, even for academic purposes is not new, as it has been frequently reported (Hewagamage and Hewagamage, 2011; Zhang et al., 2012). Authors preferred it instead of the most common standard from the Project Management Institute (PMI) –Project Management Body of Knowledge, PMBoK (PMI, 2008)- because of the students' lack of previous experience. After initial experiences (Ordieres et al., 2011), teachers found that keeping the focus at the products to be developed, instead of using an effort oriented methodology help the most to the learning process, as students always look at product level.

According to the chosen method a multiphase lifecycle is accomplished. The meaning for all the stages established by PRINCE2™ is learnt during the first three weeks of the course.

For the student's learning process, it is necessary to make clear the difference between the different roles of persons who work together on the same project, but with very different responsibilities. In order to do this, and because students from different locations –Universidad Politécnica de Madrid (UPM) and Universidad de La Rioja (UR)- and different backgrounds are involved, they are exposed to different participation experiences by playing three different roles (all of them are the available figures in PRINCE2™):

- **PM:** Project Manager, with management responsibilities. Each project is managed by a team of seven or eight PM. The number of students playing the role of PM was established according to the necessity that all students perform management tasks. It must be noted the short length of the course, just 120/150 hours of student's work (4,8/6 ECTS

assigned to the UPM and UR courses, respectively), which becomes short time considering the lack of experience of the students (Warfvinge, 2008).

- **TMg:** Team manager. A PM temporarily assigned to manage Project Engineers (TM), to produce what it was described into the Work Package document (Managing Products processes).
- **TM:** Team member, with engineering tasks development responsibilities. Each project is composed of ten or twelve TM.

The projects provided are basically oriented to learn about the project management methodology as well as to develop key competences as they include reports preparation, video presentation for the project as a commercial product as well as an individual presentation much more technical about their position in the team, the tasks carried out and their self-assessment as this tool is a beneficial for the learning process too (Crook et al., 2012). It is included the drawing preparation if required by the project topic as well as the formal budget estimation, not only for the project itself but also for its implementation.

The authors have chosen the IPMA-ICB framework (IPMA, 2006) as a reference model for competences in project management because of its flexibility and the taxonomy provided. IPMA (2006) defines the four competences studied in this work as follows:

- *Leadership* involves providing direction and motivating others in their role or task to fulfill the project's objectives.
- *Engagement & Motivation*. *Engagement* is the personal buy-in from the project manager (PM) to the project and from the people inside and associated with the project and motivation. *Motivation* of the project team depends on how well the individuals bond together and their ability to deal with both the high and low points of the project.
- *Results Orientation* means to focus the team's attention on key objectives to obtain the optimum outcome for all the parties involved. The PM has to ensure that the project results satisfy the relevant interested parties. To deliver the results required by the relevant interested parties, the PM has to find out what the different participants in the project would like to get out of it for themselves. This competence in project management behaviour is closely linked to project success
- *Teamwork* covers the management and leadership of team building, operating in teams and group dynamics.

In order to support learning and monitoring, information and communication technology (ICT) tools were used. The provided ICT environment was built by integrating some open source tools as it is described in González-Marcos *et al* (2013). Moreover, specific procedures about how to operate, how to do things, how to communicate mandatory information etc., have been developed and learning them as well as the use of the ICT system is the goal of the first module for the course, in parallel with learning about PRINCE2™.

Subsequently, students will develop a direct relationship between PRINCE2™ theory and operational procedure. This module consumes three weeks and the last activity is an individual assessment that is used as evidence for Project Management Information Systems ability and theoretical knowledge about the method of management.

During the period where the project is being developed, students realize how complex the project management becomes because different effects such as contradictions found between stakeholders, misunderstandings, time pressure, particular motivations, or lack of attention to details. Along this period the students are still learning theoretical concepts for the document structure to be delivered, legal responsibilities, as well as specific techniques useful in daily work. This procedure also shows how increasing complexity and uncertainty call for a more comprehensive inclusion of managerial and leadership knowledge respective to our teaching of advanced project management (Thomas and Mengel, 2008).

Obviously, most of the work needs to be carried out by groups or teams; however, it is based on individual knowledge. Sometimes this individual knowledge is improved because of the discussions about how to perform the work. Thus, students are responsible for their learning as well as the learning of others (Hughes, 2012).

Evaluating the learning process is an essential issue not only for students but also for teachers as they are responsible for the learning process. Unfortunately, there is no agreement on how to integrate different dimensions of learning, knowledge, skills, etc. (Huang & Yang, 2009). Therefore, authors have incorporated two different assessment methods. The first one is a formal knowledge based set of test in different period of time. The second one is a continuous assessment

oriented to estimate the project management performance and the contribution of each student to it (Qureshi et al., 2009). This project performance is based on the auditing processes carried out by the Project Board and the Owner's representatives (the teachers) as well as on the competence level gained during the daily work performed.

The auditing process has two different branches. The first one is automatically performed by a web tool developed by the authors. This application collects real-time information about project and students' progress on the Enterprise Program Management Office (ePMO) software used during the simulation. It gathers relevant information about each student's performance in their project activities (project planning activities, documents uploaded, effort allocation, use of the provided communication tools –blog, discussions, etc.). It also looks for measurable mistakes, such as the absence of relationships between tasks, the absence of links between documents and deliverables, improper effort allocations, wrong document codifications, etc. Thus, instructors are able to make periodic reports to better identify mistakes or inappropriate behaviours. In this way, the teachers can more objectively and efficiently monitor and evaluate students continuously throughout the whole course. Furthermore, students were given the right to order an on-line self-audit based on the aforementioned automatic checks.

The second branch asks for a more qualitative but still evidence-based opinion about the products being produced as well as about how the different PRINCE2's themes –risk, communication, quality and configuration- are being managed by the team. To determine the exhibited competence level the answers to different questions are gathered from different forms about products, processes and behaviours. Most of them are Likert scale based and opinions come not only from the producer of the product or responsible for the process implementation but also from different consumers of those products or participants in the process.

The numeric assessment of the different evidences considered as relevant to each competence is carried out after considering, at least, four different roles:

- The self-assessment, as it is always a relevant perception.
- The auditor
- The owner of the product being developed
- User(s) of this particular configuration item or product.

Thus, the competence assessment framework uses some kind of 360-degree overview to different activities inside the project and it collects all these evidences in a weighted integration.

4 RESULTS AND DISCUSSION

Students from two different universities (UR and UPM) were organized around eight projects (1301 to 1308). Each project was composed of seven or eight PM and ten to twelve TM. At the end of the course, more than 450 assessment forms were filled out.

The first step in any multivariate analysis is to graphically represent the individual variables using a histogram or boxplot. These graphic representations are extremely useful for detecting asymmetries, heterogeneity, outliers, etc.

In order to observe differences between perceived PM competences within each project, boxplots were used because they are a way of summarizing a distribution, take up less space than other graphical techniques and they are a quick way of examining one or more sets of data graphically (see Figure 1). The spacings between the different parts of the box help indicate the degree of dispersion (spread) and skewness in the data. A boxplot (also known as a box and whisker plot) is interpreted as follows:

- The box itself contains the middle 50% of the data. The left edge (hinge) of the box indicates the 75th percentile of the data set, and the right hinge indicates the 25th percentile. The range of the middle two quartiles is known as the inter-quartile range.
- The dot in the box indicates the median value of the data.
- The ends of the horizontal lines or "whiskers" indicate the minimum and maximum data values, unless outliers are present in which case the whiskers extend to a maximum of 1.5 times the inter-quartile range.
- The points outside the ends of the whiskers are outliers or suspected outliers.

Comparing the boxplots across groups, a simple summary is to say that the box area for one group is higher or lower than that for another group. To the extent that the boxes do not overlap, the groups are quite different from one another.

Distributions shown in Figure 1 illustrate the different opinion that each team project had about the exhibited competences of their PMs. Thus, for instance, PM team of project 1302 obtained opinions varying from 'strongly disagreement' (Likert scale

of 1) to 'strongly agreement' (Likert scale of 5), whereas the other project teams introduced better opinions on the competence level of their PMs (from 'neither agree nor disagree' –Likert scale of 3- to 'strongly agreement', with some exceptions).

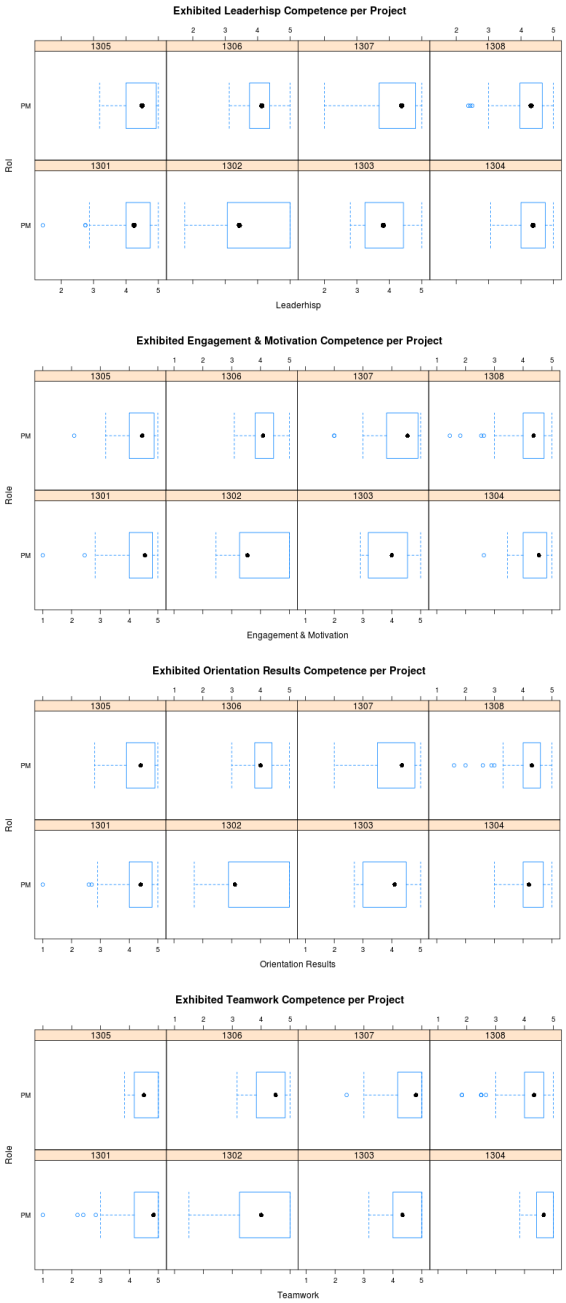


Figure 1: Competence assessment results per project.

In general, teamwork (Figure 1, bottom) was the competence with better assessments. In this case, projects 1304 and 1305 exhibited the lowest dispersions with highest assessments (between 3.8 and 5), whereas project 1302 had the highest variability in the assessments (between 1.4 and 5).

It is worth to mention that the other three competences analysed (leadership, engagement & motivation and results orientation) had very similar distributions per project. Since these three competences were evaluated at the same time by means of the same survey, it seems that the evaluation of each person was based on a global opinion that the assessor had about the assessed person without making any distinction between these competences.

The instructors team established the quality of the project (SCORE) according to the content and format of both management products (plans, business case, project reports, etc.) and specialist products (feasibility studies, engineering drawings, calculations, etc.). In this case, a 5-point Likert scale (from 1 for "very bad quality" to 5 for "very good quality") was used.

A visual inspection of all possible pairwise scatterplots in the analysed variables helps to understand the relationships between the variables. If these scatterplots are arranged in matrix format, the type of relationship existing between the pairs of variables can be understood and the outliers in the bivariate relationship identified. Such diagrams are particularly important for identifying non-linear relationships, in which case the covariant matrix may not offer a good summary of the dependence between variables (Peña, 2002).

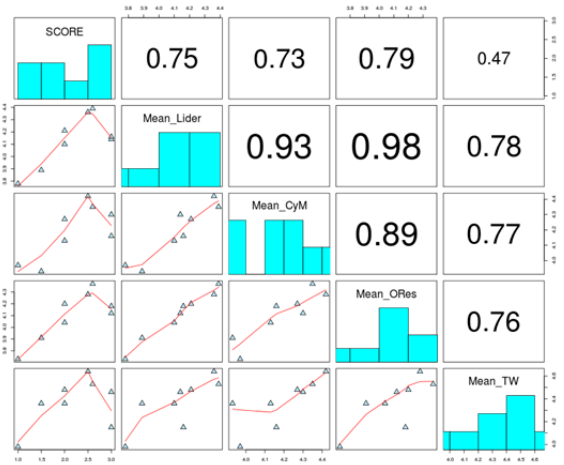


Figure 2: Relationship between project quality (SCORE) and transversal competences analysed.

Figure 2 summarizes all this information and illustrates the relationships between the project quality (SCORE) of each project and the mean value of the assessed PM's transversal competences. The

lower triangle of the matrix shows a scatterplot for each pair of variables with a polynomial approximation according to their relationship nature; the histogram of each variable appears on the diagonal; and the absolute value of the correlations with a size proportional to their magnitude is included in the top triangle. From this figure, the following conclusions can be drawn:

- The project quality (SCORE) is positively correlated with the four transversal competences analysed (Mean_Lider, Mean_CyM, Mean_Ores and Mean_TW). However, as it can be seen from the first column of the matrix, these relationships are not perfect linear (red line). Although projects 1304 and 1305 have the highest means for the four transversal competences (points located at the top of each scatterplot), projects 1307 and 1308 obtained the highest SCORES (points located at the right of each scatterplot). That is, large doses of leadership, engagement & motivation, results orientation and teamwork are important to ensure high quality of the project, but they are not the only relevant variables.
- The highest correlation between the project quality (SCORE) and any of the PM's transversal competences analysed is found for the competence named results orientation (Mean_ORes). This result is consistent with the importance that the International Project Management Association (IPMA) gives to this competence: results orientation in project management behaviour is closely linked to project success (IPMA, 2006). In summary, higher quality projects were attained by PM teams that were able to develop project teams focused on results in changing environments.
- Project quality (SCORE) and teamwork (Mean_TW) have the lowest correlation (0.47).
- Although there is a high correlation between the four transversal competences analysed, the strongest relationship is found between leadership (Mean_Lider), engagement & motivation (Mean_CyM) and results orientation (Mean_ORes). This result is consistent with our previous observation related to the evaluation of these competences through the same survey: each assessor evaluated these competences to each person without making any distinction between.

By defining the effort ratio as the relationship between the total number of hours claimed by the whole project team and the total number of planned hours, it is observed a strong correlation (0.83) between this variable and the project quality (see Fig. 3). This result suggests that the lower deviations between planned and actual activity of the project team, the higher final quality of the project.

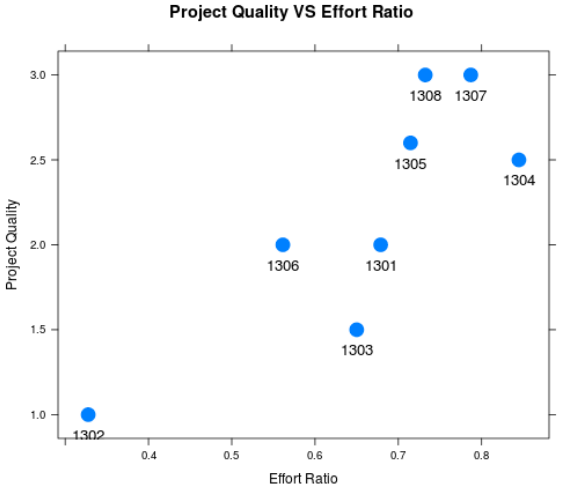


Figure 3: Relationship between project quality (SCORE) and effort ratio.

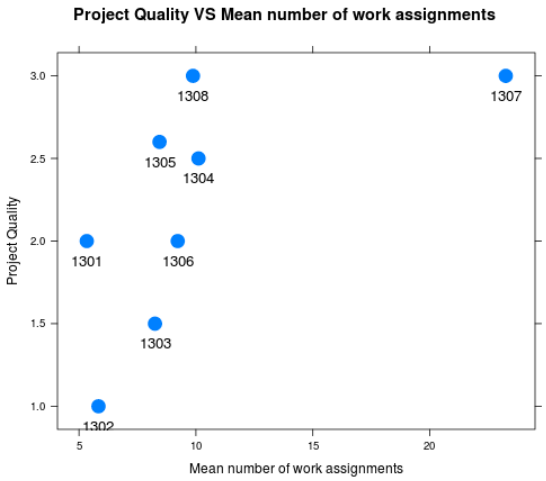


Figure 4: Relationship between project quality (SCORE) and results orientation competence.

Fig. 4 illustrates the relationship between the mean number of work assignments and the final project quality. In this case, the correlation between the two variables analysed is lower. That is, a higher number of work assignments do not necessary mean a better project quality. For example, although the number of work assignments defined in the project 1308 was half of the work assignments defined in the project 1307 both projects had a similar project quality.

Regarding the teamwork competence (Fig. 5 and Fig. 6), it is highly related to both the effort ratio

(correlation equal to 0.78) and the size of the project team (correlation equal to 0.85). These results illustrate how the teamwork feeling increases as the estimated effort was close to the actual effort (Fig. 5). The same feeling was identified taking into account the size of the project team (Fig. 6).

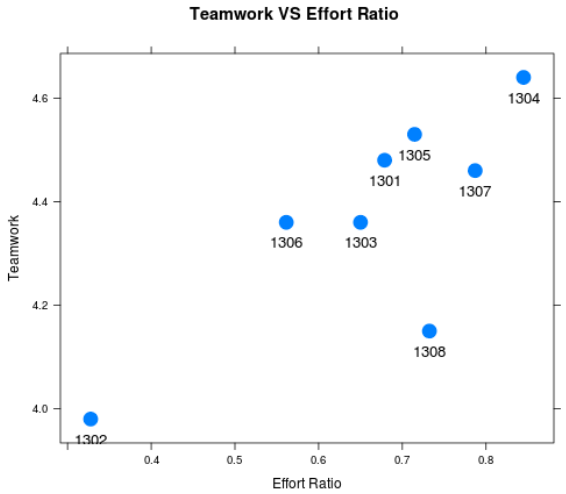


Figure 5: Relationship between teamwork competence and effort ratio.

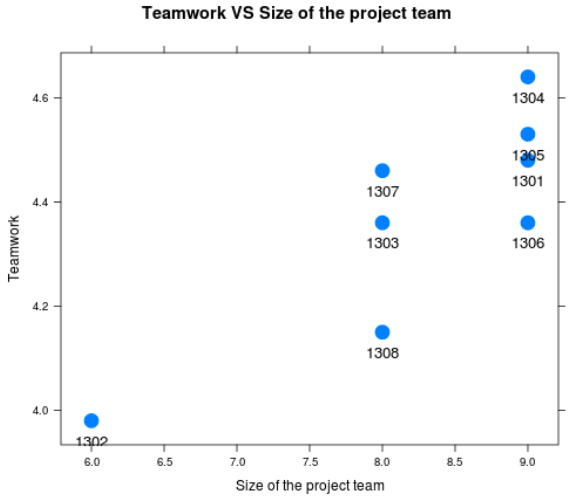


Figure 6: Relationship between teamwork competence and size of the project team.

5 CONCLUSIONS

This paper has presented an integrated framework that allows competence assessment in project management by putting it into practice in a simulation environment. The system is distinguished by the assessment of individual competences by means of a set of performance indicators. The indicators are obtained from both analytical evidence and the opinions of other participants in the simulation in relation to the skills demonstrated by the candidate's specific actions. Furthermore, the system allows obtaining information about the level of demonstrated competences not by the measurement of individual knowledge, but as the result of their use in a simulation environment.

An analysis of the PM's exhibited competences demonstrated the relationship between some transversal ('human') competences and the project success. In the academic course analysed, the strongest relationship was found between results orientation and project success, whereas teamwork did not seem to be correlated with the project quality. The authors attribute the lowest relationship between teamwork and project success to the lower number of pieces of evidence used to assess this competence.

Another interesting result is the strong relationship found between leadership, engagement & motivation and results orientation. Taking into account that these competences were evaluated at the same time by means of the same form, it seems that each assessor used a global opinion about the assessed person to fill out the form. That is, no distinction between these competences was made.

For the future authors look to improve the assessment of these competences by using more pieces of evidence. Furthermore, authors will extend the number of competences assessed by the presented framework as well as to use the collected data for early detection of problems inside the project, and to improve the learning procedure by means of the gathered data.

ACKNOWLEDGEMENTS

The authors wish to recognize the support of the "Vicerrectorado de Profesorado, Planificación e Innovación Docente" of the University of La Rioja, through the "Dirección Académica de Formación e Innovación Docente" as part of this work was supported by its grant APIDUR 2013.

REFERENCES

Abernethy, K., Piegari, G., Reichgelt, H., 2007. Teaching project management: an experiential approach, *Journal of Computing Sciences in Colleges*, 22(3), 198-205.

APM, 2008. *APM Competence Framework*. Association for Project Management.

Barron, S., 2005. Assessing project management learning – how can it make a difference? In *2nd Project Management Conference: Excellence in Teaching, Learning and Assessment*. Bournemouth, Sep. 15–16.

Beltrán, J. A., 1999. Aprender en la universidad, en Ruiz, J. (Coord.). *Aprender y enseñar en la universidad. Iniciación a la docencia universitaria*. Jaén: Universidad de Jaén.

Berggren, C., Söderlund, J., 2008. Rethinking project management education: Social twists and knowledge co-production. *International Journal of Project Management*, 26, 286-296.

Clark, R., 2008. Project Management: The Key to Engineering Employability. In *SEFI - 36th Annual Conference*. Aalborg, Denmark, 2 - 5 July.

Council of the European Union, 2004. "*Education and Training 2010". The Success of the Lisbon Strategy Hinges on Urgent Reforms* [Online]. Council document 6905/04 of 03 march 2004. Available: http://ec.europa.eu/education/policies/2010/doc/jir_council_final.pdf.

Crespo, E., Alba-Elías, F., González-Marcos, A., Ordieres, J. and Castejón-Limas, M., 2011. Project management learning in a collaborative distant learning context –an actual on-going experience. *In 3rd International Conference on Computer Supported Education*, pages 105–110.

Crook, A., Mauchline, A., Maw, S., Lawson, C., Drinkwater, R., Lundqvist, K., Orsmond, P., Gomez, S., Park, J., 2012. The use of video technology for providing feedback to students: Can it enhance the feedback experience for staff and students? *Computers and Education, 58* (1), 386-396.

European Commision, 2004. *Implementacion of "Education and Training 2010". Key Competences* [Online]. Available: http://ec.europa.eu/education/policies/2010/doc/basicframe.pdf.

González-Marcos, A., Alba-Elías, F. and Ordieres-Meré, J., 2013. A Web-based Support System for Providing Effective Monitoring, Feedback and Evaluation in Project Management Education. *In 5th International Conference on Computer Supported Education*, pages 254–257.

Hewagamage, C.; Hewagamage, K., 2011. Redesigned framework and approach for it project management. *International Journal of Software Engineering and its Applications*, 5 (3), 89–106.

Huang, S., Yang, C., 2009. Designing a semantic bliki system to support different types of knowledge and adaptive learning. *Computers & Education, 53* (3), 701-712.

Hughes, J., 2012. The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. *Journal of Technology and Teacher Education*, 277-302.

IPMA, 2006. *ICB: IPMA Competence Baseline*. Version 3.0. International Project Management Association.

Navaridas, F., 2004. *Estrategias didácticas en el aula universitaria*. Logroño: Universidad de La Rioja.

Office of Government Commerce, 2009. *Managing Successful Projects with PRINCE2(TM)*. Office Of Government Commerce.

Ojiako, U., Ashleigh, M., Chipulu, M., Maguire, S., 2011. Learning and teaching challenges in project management. *International Journal of Project Management*, 29, 268–278.

Ordieres-Meré, J., González-Marcos, A., Alba-Elías, F., 2011. The use of computers for graduate education in project management. Improving the integration to the industry. *In the 2011 International Conference on Frontiers in Education: Computer Science & Computer Engineering*, pages 233–238.

Pant, I., Baroudi, B., 2008. Project management education: the human skills imperative, *International Journal of Project Management*, 26, 124-128.

Peña, D., 2002, *Análisis de datos multivariantes*. McGraw Hill.

PMI, 2002. *Project Manager Competency Development (PMCD) Framework*. Project Management Institute.

PMI, 2008. *A Guide to the Project Management Body of Knowledge* (PMBOK Guide), 4th Edition. Project Management Institute.

Qureshi, T. M., Warraich, A. S., Hijazi, S. T., 2009. Significance of project management performance assessment (PMPA) model. *International Journal of Project Management*, 27 (4), 378-388.

Reif, H. L., Mitri, M., 2005. How university professors teach project management for information systems. *Commun. ACM*, 48(8), 134-136.

Sense, A., 2007. Structuring the project environment for learning, *International Journal of Project Management*, 25, 405-412.

Thomas, J., Mengel, T., 2008. Preparing project managers to deal with complexity –Advanced project management education. *International Journal of Project Managemen*t, 26 (3), 304.

Warfvinge, P., 2008. A generic method for distribution and transfer of ECTS and other norm-referenced grades within student cohorts. *European Journal of Engineering Education*. 33 (4), 453-462.

Zhang, L., He, J., Zhang, X., 2012. The project management maturity model and application based on PRINCE2. *Procedia Engineering*, 29, 3691–3697.

Getting it Fast
An Information Systems Oriented Semantic Web Curriculum

Daniela Giordano and Francesco Maiorana
Department of Electrical, Electronic and Computer Engineering, University of Catania, Viale A. Doria, 6, Catania, Italy
{daniela.giordano, francesco.maiorana}@dieei.unict.it

Keywords: Information Systems Applications, Mash-ups, RDF, RDS, OWL, Linked Open Data, Knowledge Representation and Reasoning, Ontology Engineering, Triplify, SPARQL, Gruff, Jena.

Abstract: The fast growing importance of the semantic web and semantic web applications is demonstrated by the exponentially growing amount of semantic data produced on the web and by the rise of the linked open data movement. Besides this strong interest there is a request both from academia and industry for well-prepared students in order to minimize the training effort needed to prepare them for productive work. This paper describes the teaching experience of a Master's course entitled "Laboratory in software design and development – semantic technologies". A detailed description of the curriculum, the rationale underlying the choice of content and the software tools used, as well as the main lessons learnt from the experience, are presented.

1 INTRODUCTION

The enormous impact of the World Wide Web (WWW) is hindered by some problems related to:

- data representation: all the data on the WWW are semantics-free. This gives rise to disaligned, inconsistent and unrelated information.
- entity update: an update on an entity does not affect all the pages that semantically reference it.

The semantic web is able to overcome these difficulties and is gaining momentum day by day. The rise of the semantic web has created a strong demand by industry for people trained in semantic technologies as well as ontologies. This demand is estimated to quickly grow in the following years. Paralleling this strong interest there is a flourishing growth of courses in academia and professional institutions; however, these courses are often single courses within larger programs. The authors in (Neuhaus, 2011) identified only one complete academic program completely devoted to education in applied ontology (a Master's program at the University of Buffalo) and 21 programs that offer ontology-centered topics.

Applied ontology and the semantic web are relatively young disciplines and therefore there is not yet a general consensus on questions related to several aspects ranging from the curriculum to even the terminology.

In the ACM/IEEE (ACM/IEEE, 2013) computer science curriculum, for example, the knowledge required for an ontologist is spread among courses or topics within courses such as the elective "Advanced Representation and Reasoning" in the field of Intelligent Systems (IS), dealing with ontology engineering and, in the Computational Science field, the elective course entitled "Data, Information and Knowledge", dealing with "Knowledge: ontologies, triple stores, semantic networks, rules", and so on. First, this paper will briefly review teaching literature, curriculum examples and didactic tools in the main areas of the course: ontology design, ontology query languages, ontology programming framework, mash-up and semantic web applications. The design of an introductory lab in semantic web technologies, worth 3 ECTS, will also be presented.

Among the several collections of semantic web educational resources to which the reader can refer (Diederich, 2007), there exists a repository built by the European Association for Semantic Web Education (Diederich, 2006). In spite of this field being relatively young, there are outstanding experiences in teaching ontology engineering and ontology design. In (Rewctor, 2004) the authors report the results of a several-years-long experience

in teaching the OWL language and ontology design, along with common errors and common design patterns. Several ontologies have been used as didactic examples to teach ontology design and OWL, such as Pizza, Wines and Marsupials.

Recently, a didactic example in the biological domain was proposed by (Schober, 2012), where the authors propose an OWL ontology describing zoo animals. The ontology, with respect to other large biomedical ontologies, has the advantage of massively reducing the number of classes and overall complexity, allowing for fast memorization and orientation, yet it uses all the major ontology constructs, with related modelling challenges together with some Ontology Design Pattern (ODP). The proposed ontology uses a light version of Bitop (Beißwanger, 2008) as a top level ontology. Moreover, the ontology covers a domain that involves common knowledge and therefore can be easily understood by the majority of people, especially in the life science domain. The ontology was used in a complete curriculum for ontology design described in (Boeker, 2012) where the authors present a teaching experience of a one-week workshop organized around 16 modules, addressing 4 main themes with an increasing level of design complexity: basic principles, practical ontology design, using top-level ontologies and ODP.

Various works address the development of software tools used in training students, training and teaching ontology query languages, and examples of clear and didactically sound tutorials are also available. A recent example of a software training tool is described in (Gerber, 2010) where the authors present a web-based SPARQL trainer that allows the tutor to design a course along a set of concepts that are to be tested. The tutor prepares a set of questions and a dataset. Each question is provided with a query solution against which the student solution is compared. The comparison is based on the result sets: the query solution and the student solution result set must coincide both in the elements and in their order.

For mash-up and semantic web applications, a good tutorial can be found in (Della Valle, 2008) where the author presents a semantic web application that expects a music style as an input, retrieves data from online music archives and event databases, merges them by a bridge ontology and allows the user to explore events related to artists that practice the required style.

The aim of this paper is to give a detailed overview of the curriculum taught in a Master's-level class in Informatics Engineering in order to "describe in more detail" how some of the knowledge and skills indicated in (Neuhaus, 2011) can be acquired. Our curriculum is "IT-oriented" with a particular emphasis placed on the knowledge and deployment of "IT systems involving many components in addition to the ontology itself" (Neuhaus, 2011). This type of curriculum fits well with the knowledge and background of Informatics Engineering and Computer Science students and provides a means of broadening both data and knowledge design capabilities, as well as system integration and software development skills.

The paper is organized as follows: section 2 presents the teaching context and the contents; section 3 describes the lessons learnt, highlighting strengths and weaknesses of the approach used in the course; section 4 reports conclusions and describes further work.

2 TEACHING CONTEXT AND CONTENTS

The teaching experience pertains to a 30-hour course within the Informatics Engineering Master's degree at the University of Catania. The course was taught in the 2011 Fall term. The course was an elective one, leading to a final certification. It was attended by both first year and second year students. Management of the class was not particularly easy due to the students' varying levels of knowledge. The course schedule was a three hour meeting per week for ten weeks. The course was designed around semantic technologies, with the aim of demonstrating during the classes the main technologies of the semantic web and the main tools used to design and implement semantic applications. As a university rule, the course was graded on a pass/fail basis.

The students attending the course could get acquainted with the following topics and questions:

1) Data transformation from a relational database to RDF. After a review of the main tools, Triplify (http://triplify.org/Overview) was chosen due to its simplicity and its tight relationship with database technologies and the PHP web programming language, tools that are well-known to all the students attending the course.

2) Querying triple stores/RDF files. The main tool used was the SPARQL query language. The Gruff browser was used both in a standalone version and with the AllegroGraph server. Among the different advantages of the Gruff browser, the graph

representation capabilities were considered a definitive plus.

3) Relations and ontology design using RDFS and OWL. Protégé was the tool chosen, due to its flexibility, integration with reasoners, rich availability of plugins and its free availability.

4) Linked Open Data: design and implementation of a mash-up from linked data. The Application Framework, Jena, was chosen.

The course also offered students possibilities to familiarize themselves with the following core skills and knowledge:

- Clarifying the purpose of a given ontology
- Judging what kinds of ontologies are useful for a given problem
- Identifying, evaluating and using software tools that support ontology development
- Using (reading and writing) different representation languages
- Conducting ontological analyses
- Using a modern programming language
- Working in teams

The class activities were supported by an on-line class site developed with the Moodle platform, used for sharing lecture notes, projects and communications and greatly simplified the management of class activities.

The class schedule is reported in Table 1 with an indication of the time required for each topic.

The table gives an overview of the curriculum content that was organized into 4 main parts or modules, with an increasing level of difficulty and with a broader view of the covered topics. Each module was further divided into 2 or 3 units. The modularization of the curriculum allows for an easier customization for other courses and classes with different backgrounds.

The first module presents the Semantic Web in comparison to the WWW and its main contributions. The RDF language and RDF graph are presented along with the major serialization languages: N-triples, turtle and XML. Concepts like reification, blank node and list and their impact in designing are presented and applied to examples and exercises. An introduction to the main ontological languages RDFS, RDFS-Plus, OWL and their expressiveness is undertaken.

The second unit deals with data transformation and query techniques. The data transformation technique was appreciated by the students and served as a common thread between their strong background on database and web programming and the course content. After a review of the main tools, such as D2RQ, Triplify was chosen due to its simplicity and its affinity with SQL and the PHP development environment. This allowed students to publish RDF triples from their own databases as well as from data present in Bulletin Board, Content Management systems and so on. The students

Table 1: Class schedule.

Topic	#H	Tools	Readings	Assignment
Part 1: Basic Principles				
The Semantic web	3		(Allemang, 2011) Ch. 1	
RDF and RDF graph. Serialization languages.	3		(Allemang, 2011) Ch. 2-3	
Reification, blank node, list in RDF. RDFS, RDFS-Plus, OWL	3		(Allemang, 2011) Ch. 3-4	
Part 2: Data transformation and querying				
Data transformation tools. Lab session: use of Triplify	3	Triplify Allegro-graph	Auer, 2009	Triplify a Database
SPARQL and SPARQL 1.1	4	Allegro-graph/ Gruff	Ch. 5 (Allemang, 2011), Ch. 6 (Liyang, 2011), (Bizer, 2009)	Design a set of queries and evaluate their performance
Part 3: Ontology design				
Ontology engineering. Use of Protegè	4	Protegè/ OWLViz	Ch. 4 (Della Valle, 2009), Ch. 6, 7 8, 9, 11, 12 (Allemang, 2011) ch. 4, 5, 12 (Liyang, 2011)	Design an ontology
The Jena framework	4	Eclips/ Jena/ MYSQL	Ch. 12, 13, 14, 15 (Liyang, 2011), (McBride, 2002), (Carroll, 2004)	Develop an application that integrates data from FOAF and DBpedia
Part 4: Semantic web applications: mash-up, linked data and visualization				
Visualization framework: Exhibit	2	Exhibit	Huynh, 2007	Visualize events data on google maps
A semantic web application: mash-up and linked data	4	Combi-nation of tools	(Bizer, 2009 (b)) and (Heath, 2011), Della Valle, 2008) Ch. (Della Valle, 2009)	Develop a mash-up application (optional)
Total	30			

became acquainted with large RDF files and began to design RDF graphs starting from an E/R schema.

The original work by (Auer, 2009) describing the design of Triplify was used as a main reference. The best works were presented by the students, which benefited the entire class. The use of a shared pool of projects, (Giordano, 2004) can be viable, even if course is in its first edition, by using the projects developed throughout the year, to facilitate the students in overcoming difficulties, finding technical details and solutions as well as to gain inspiration from the work of their peers.

The subsequent step was the introduction of the SPARQL language to query a RDF triple store. The SPARQL language was selected since it resembles the well-known database query language SQL. The use of triples, graph patterns and other features allowed the students to practice with the underlying RDF graph and with serialization languages. Moreover, the students were free to use OWL and domain ontologies and to practice with large triple stores such as DBpedia. For all these reasons the introduction of SPARQL was considered a beneficial step in this phase of the course. The students were free to choose the RDF store where they published their data. The suggested choice was the AllegroGraph RDF store (http://www.franz.com/agraph/allegrograph/) and the Gruff graph based triple store browser (http://www.franz.com/agraph/gruff/) due to the graph representation of the results set. Other common choices were the Sesame (http://www.openrdf.org/) and Joseky (http://joseki.sourceforge.net/) RDF stores.

Chapter 5 of (Allemang, 2011) and chapter 6 of (Liyang, 2011) were used as main references. All the types of queries were covered. The unit started with the select query form presenting triple and graph patterns, query modifiers, optional patterns that could be nested, filter conditions on different types and regular expressions, union of graph patterns, background and named graphs to query multiple graphs. The other types of query forms, namely construct and rules, describe and ask, were then covered followed by the new SPARQL 1.1 features such as aggregate functions, group by, sample and bound, subqueries, negation, expressions with SELECT, property paths, transitive queries, federated queries and the SPARQL update 1.1 standard. In order to make the students aware of performance issues in query formulation the problem of benchmarking was presented through studies such as (Bizer, 2009) and (Schmidt, 2009). These works were also used as query examples. Tools such as (http://ftp.heanet.ie/disk1/download.sourceforge.net/pub/sourceforge/b/project/bs/bsbmtools/bsbmtools/bsbmtools-0.2/) were used by the students to profile queries and experiment with different query formulations. The students experimented with different datasets such as the one suggested in (Bizer, 2009), FOAF RDF files, well-known public domain ontologies such as the Wine ontology and so on. Figure 1 shows an example of a graph created with the Gruff browser and Table 2 shows a sample output of the bsbmtools available in the Berlin SPARQL Benchmark.

The third part of the course was divided among two related units. The first dealt with ontology design principles. The main references were chapter 4 of (Della Valle, 2009) as well as chapters 6, 7 for RDFS design and chapters 8, 9, 11, 12 from (Allemang, 2011) for OWL-related design. Finally, chapter 4 for RDFS design, chapter 5 for OWL design and chapter 12 for tools and resources from (Liyang, 2011) were also suggested to the students for reference. Suggestions on how to present the material are reported in the Lessons Learnt section. As a working tool, Protégé was chosen (http://protege.stanford.edu/) with the OWLViz (http://www.co-ode.org/downloads/owlviz/) for class visualization. The approach was to follow the Protegè tutorial (Horridge, 2009) as well as sharing suggestions such as those reported in (Rewctor, 2004) and (Noy, 2004), just to cite some examples. An initial methodology and a practical example for converting an E/R model into an OWL model was presented, along with some ODP. Other literature was suggested as further readings, i.e., (Poveda-Villalón, 2010) and (Hammar, 2010). Two practical ontology design exercises were also proposed, the first one about the design of an educational portfolio and the second one regarding the sharing of learning objects. As a framework to develop semantic applications and manage ontologies, the Jena framework was chosen due to the large use of Java technologies in both academia and the professional world, as well as its free availability.

The Eclipse IDE and the MYSQL database to store persistent models, were also suggested for data manipulation in a database. As a reference for practical applications, chapters 12, 13, 14 and 15 of (Liyang, 2011) were suggested.

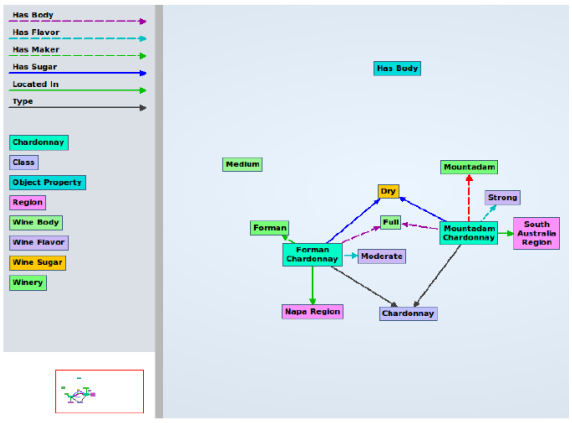
To present the Jena framework and its main functionality and usage, the guidelines of presentation indicated in (McBride, 2002) and (Carroll, 2004) were followed. Following this in-depth introduction on the architecture and main features, practical examples on creating, using, manipulating and navigating through RDF graph,

reading and storing to file as well as to database were given with practical applications and working code. Practical examples performing queries along with using reasoners were also presented. The tutorial used in the previous model was recreated using, programmatically, the OWL API. Finally, a complete example of a linked data application in the style of (Liyang, 2011) was illustrated. The example was used as a starting point for a linked data application. The practical approach to linked data was followed by a theoretical overview. The books (Bizer, 2009 (b)) and (Heath, 2011) were used as reference material and an overview of the main concepts, design principles, techniques and recipes for publishing and consuming linked data was presented to the students.

```
PREFIX vin: <http://www.w3.org/TR/2003/CR-owl-guide-20030818/wine#>

SELECT * WHERE {
  ?winery a vin:Chardonnay .
  MINUS{?winery vin:hasBody vin:Medium .}
  ?winery ?property ?value
}
```

(a)



(b)

Figure 1: An example of a SPARQL 1.1 query (a) with the graph results visualized with the Gruff browser (b).

Finally, the fourth part completed the picture by presenting visualization tools and strategies to design and implement a mash-up semantic application. Exhibit was selected as the visualization tool due to is simplicity and flexibility in addition to its free availability (Huynh, 2007) together with tutorials and other learning material. In particular, Exhibit 3.0 was presented to the students, together with its use in a local environment.

In the original plan, all the technologies tools and techniques had to converge in the presentation of the final project. The aim was to present a mash-up application that integrates different sources of data in different formats, (e.g., relational, XML and RDF) by designing a bridge ontology integrating all these data in a unifying view. The reference material was the tutorial (Della Valle, 2008) as well as chapter 9 of (Della Valle, 2009) where the design principles and main techniques and implementation details are presented. This kind of application should be considered as the concluding and unifying aspect of the course, where all the principles and technologies can be applied to build a working semantic application.

Table 2: Result of the performance profile of the query on the dataset described in (Schmidt, 2009) with 1.000 as scale factor. The scale factor used is related to the number of products.

Query: Find information of the Person whose label is "reexhibit"
describe ?x
where {
?x rdfs:label "reexhibit".
}

Output of the bsbmtools-0.2 Scale factor: 1000 Number of warmup runs: 20 Seed: 808080 Number of query mix runs (without warmups): 50 times min/max Querymix runtime: 2.6075s / 3.1330s Total runtime: 140.452 seconds QMpH: 1281.58 query mixes per hour CQET: 2.80904 seconds average runtime of query mix CQET (geom.): 2.80503 seconds geometric mean runtime of query mix Metrics for Query: 1 Count: 50 times executed in whole run AQET: 0.004793 seconds (arithmetic mean) AQET(geom.): 0.004764 seconds (geometric mean) QPS: 208.64 Queries per second minQET/maxQET: 0.00429478s / 0.00822679s Average result (Bytes): 1320.00 min/max result (Bytes): 1320 / 1320 Number of timeouts: 0

The guiding factors for choosing the tools were: the efficiency, the ease of use and its connection with the database technologies for Triplify, the visualization aids for Gruff, the free and open source availability together with its large use in ontology development for Jena, and the relative ease of use and rich set of graphical features for Exhibit.

Overall there was a positive appreciation of the course, particularly from the more motivated students. Some of them expanded upon the course content with a Master's thesis.

3 LESSONS LEARNT

Analyzing the teaching experience, it is possible to make the following considerations. The initial module on the semantic web, although an interesting and necessary step that gives an overview of the potentiality and breadth of the fields, can be reduced to a minimum and allow the important concepts to emerge from practical case studies such as querying an RDF graph and ontology design.

The use of Triplify was a good ice breaking activity allowing students to start from what they already knew well, namely database design and web programming, bringing them to fast production of their RDF data and hence reading and navigating the data produced.

The use of SPARQL for querying RDF store was a definitive plus, serving as a bridge between different concepts: it allows the student to master serialization languages, ontology representation and navigation, triple and graph pattern, query language, and so on. The use of a benchmarking tool can be of great benefit by allowing the students to place more attention on query design and underlying execution mechanisms. This type of knowledge can help in obtaining a better understaning of both the querying and, more broadly, of the ontology design processes.

The ontology design step is worth further consideration. This can be done by allowing more time for student projects and by sharing both their projects and the peer reviews along with the trainer's feedback and comments (Giordano, 2004). The use of conversion from E/R to OWL is a good starting point. Supplementing ODP by applying it to a real case scenario could be of great benefit. The amount of material suggested to the student should be covered inside the development of a practical project in order to avoid an overwhelming effect.

The programming side of the linked data and mash-up applications can be the central aspect of courses for Informatics Engineering and Computer Science students. By leveraging their strong background on software design and programming, by scheduling the necessary time and granting a valid number of credit units, there is the possibility to develop, through group projects, interesting examples of applications.

Some students reported the preference to begin with ontology design and then view the querying aspects. The best curriculum may then be designed around a parallel development of query, ontology design and programming framework to converge as soon as possible with the design and implementation of linked data and mash-up in a semantic web application.

Due to the breadth and richness of the field along with the cognitive load needed for mastering all the topics and theoretical concepts as well as the set of technical expertise, both as an IS engineer and as designer and programmer, a course with more than 30 hours is required. In practice, due to an initial underestimation of the overall workload, the requirement for the final project was turned into an optional one (only 10% of the students completed it) and the course assessment was performed based on three comprehensive exercises that demonstrated a working knowledge of the techniques and tools showcased during the course. With a course of 40 to 60 hours, more time could be devoted to practical projects with a wider scope, to be developed both in lab sessions, under the guidance of a tutor, and at home, leaving the contents as described. Some informal feedback collected by students who entered the workforce after graduating and were assigned responsibilities involving semantic web technologies, commented that although the contents were covered in a very condensed timespan, they provided an effective groundwork.

4 CONCLUSIONS AND FURTHER WORK

This paper has presented a curriculum, oriented towards Information Technology, suited for a course on semantic web and ontologies. The syllabus is organized around 4 major modules allowing for easy customization. The curriculum, by leveraging on the database, software engineering and web programming backgrounds of the participants, lays the foundation for mastering semantic web theoretical essentials, as well as the tools and techniques necessary to develop applications that exploit the full potential of this emerging field.

As further study a systematic analysis of the mini-projects handed in by the students is planned, to document common errors and pitfalls, for example, in the query formulation with respect to both the desired result and the execution efficiency, to adjust accordingly further versions of this elective course.

REFERENCES

Allemang, D., Hendler, J., 2011. *Semantic web for the working ontologist. Effective Modelingin RDFS and*

OWL. Second Edition.(2011) Morgan Kaufmann.

Association for Computing Machinery, IEEE-Computer Society, 2012. Draft Computer Science Curricula 2013.

Auer, S., S. Dietzold, Lehmann, J., Hellmann, S., Aumueller, D., 2009. Triplify - Light-weight Linked Data publication from relational databases. Proceedings of the *World Wide Web Conference* (WWW) (2009), pp. 621-630.

Beißwanger, E et al., 2008. BioTop: An Upper Domain Ontology for the Life Sciences - A Description of its Current Structure, Contents, and Interfaces to OBO Ontologies, In: *Applied Ontology* 3:4, pp. 205- 212.

Bizer, C.,Schultz, A., 2009. The Berlin SPARQL benchmark. *International Journal on Semantic Web and Information Systems* 5(2): pp. 1-24.

Bizer, C., Heath, T., Berners-Lee. T., (2009 b) "Linked data-the story so far." *International Journal on Semantic Web and Information Systems (IJSWIS)* 5.3 (2009): 1-22.

Boeker, M , Schober, D., Raufie, D.,Grewe, N., Rohl, J., Jansen, L., Schulz, S., 2012. Teaching good biomedical ontology design. In Cornet R & Stevens R (eds.), Proceedings of the *3rd International Conference on Biomedical Ontology* (ICBO 2012), KR-MED Series, July 21-25, 2012, Graz, Austria. URL http://ceur-ws.org/Vol-897/sessionJ-paper25.pdf.

Carroll, J. J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K., 2004. Jena: implementing the semantic web recommendations. In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters,* pp. 74-83. ACM.

Della Valle, E., Celino, I., Cerizza, D., 2009. *Semantic Web: dai fondamenti alla realizzazione di un'applicazione*. Pearson – Addison Wesley.

Della Valle, E. 2008. RSWA 2008. Realizing a Semantic Web Application. Tutorial presented at the 7th *International Semantic Web Conference* (ISWC 2008), Karlsruhe - Germany.

Diederich, J., M. Dzbor, Maynard, D., 2007. REASE— The repository for learning units about the Semantic Web. *New Review of Hypermedia and Multimedia* 13(2): pp. 211-237.

Diederich, J., Nejdl, W., Tolksdorf, R., 2006. EASE: The European Association for SemanticWeb Education. In *Proceedings of the Semantic Web Education and Training Workshop (SWET'06), co-located with the First Asian Semantic Web Conference (ASWC 2006)*. China: Beijing.

Gerber, D., Frommhold, M., Martin, M., Tramp, S., Auer, S., 2010. Learning semantic web technologies with the web-based SPARQLTrainer. In *Proceedings of the 6th International Conference on Semantic Systems* pp. 26. ACM.

Giordano, D., 2004. Shared values as anchors of a learning community: A case study in information systems design. *Journal of Educational Media*, 29, 3, 213-227.

Hammar, K., and Sandkuhl, K., 2010. The State of Ontology Pattern Research - A Systematic Review of ISWC, ESWC and ASWC 2005–2009. In: Blomqvist, E., Chaudhri, V., Corcho, O., Presutti, V., Sandkuhl, K. (eds.) Proceedings of the *2nd International Workshop on Ontology Patterns* - WOP2010. CEUR-WS.org, Shanghai, China (2010).

Heath, T., Bizer, C., 2011. *Linked data: evolving the web into a global data space*. Morgan & Claypool Synthesis Lectures.

Horridge, M., 2009. A practical guide to building OWL Ontologies Using Protegé 4 and CO-ODE tools Edition 1.3 available at: http:// owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/ retrieved September 2013.

Huynh, D. F., Karger, D. R., Miller, R. C., 2007. Exhibit: Lightweight Structured Data Publishing..

Liyang, Y. 2011. *A Developer's Guide to the Semantic Web*. Springer.

McBride, B., 2002. Jena: A Semantic Web Toolkit. *IEEE Internet Computing* 6(6): pp. 55-58.

Neuhaus, F., Florescu, E., Galton, A., Gruninger, M., Guariono, N., Obrst, L., Sanchez, A., Vizedom, A., Yim, P., Smith, B., 2011. Creating the ontologist of the future. *Applied Ontologist*, 6, pp. 91-98.

Noy, N., Mc Guinness, D. L., 2004. Ontology development 101: A guide to creating your first ontology. Stanford University.

Poveda-Villalón M., Suárez-Figueroa M.C., and Gómez-Pérez A. 2010. A double classification of common pitfalls in ontologies. *Workshop on Ontology Quality (OntoQual 2010)*, Co-located with EKAW 2010, October 15, 2010, Lisbon, Portugal.

Rewctor, A., Drummond, N., Horridge, M., Rogers, J., 2004. OWL pizzas: practical experience of teaching OWL-DL: common errors & common patterns. Knowledge in the age, Springer.

Schmidt, M., Hornung, T., Lausen, G., Pinkel, C., 2009. SP2Bench: A SPARQL Performance Benchmark.

Schober, D., Grewe, N., Rohl, J., and Boeker, M. 2012. ZooAnimals.owl: a didactically sound example-ontology for teaching description logics in OWL 2. *OBML 2012 Workshop Proceedings*.

Teaching Web Programming
An Approach Rooted in Database Principles

Francesco Maiorana
Department of Electrical, Electronic and Computer Engineering, University of Catania, Viale A. Doria, 6, Catania, Italy
francesco.maiorana@dieei.unict.it

Keywords: Web Programming, Database Driven Applications, Security, Transaction, Stored Procedure.

Abstract: The World Wide Web has become a medium not only for delivering information but also for providing applications that can be used by a variety of instruments, including mobile devices. For these reasons the importance of teaching these skills as well as the difficulties of organizing a vast amount of concepts and technologies in a fast-paced environment driven by strong industrial interest and competition are widely recognized. This work discusses a case study on teaching web programming, presenting an approach suitable for both high school and university courses. The main innovation of the curriculum is the content presentation, which focuses on small incremental steps, spanning the majority of term lab sessions, as well as a teaching a case study which centres on user login management. The case study covers all the major topics in web programming and teaches students to develop, as well as apply and integrate, major database concepts such as transactions, stored procedures and security issues. This lays the foundation for developing a larger project, which is suggested as the final step of the curriculum.

1 INTRODUCTION

According to the National Science Foundation (2012), (Bureau of Labor Statistics), the occupational outlook for Information Security Analysts, Web Developers, and Computer Network Architects between 2010-2020 is expected to increase at a rate of 22%, which is faster than average. The strong interest is confirmed by the increase in web technologies. In (Gundotra, 2009), for example, during his keynote speech at the Google I/O 2009 conference, the Vice President of Engineering at Google stated that, "The web has won. It has become the dominant programming model of our time".

At the same time, Academia has given web development an increasing level of importance, reflected by the large number of related courses, from entry level to more advanced. Also in research there is a flourishing of conferences and special issue journals such as (Connolly and Miller, 2013) with "two inter-related aims. The first of these is to communicate current approaches to teaching web development. The second is to reignite the conversation about what role, if any, that web development should have in current computing curricula". The same trend is also evident in K-12 education, where web development is recognized as an important topic in all recently revised curriculum. These include (CSTA 2011), (Computing at Schools, 2012), (New Zealand Ministry of Education 2010), as well as new curricula designed at the national level, such as (Astrachan, 2012) or initiatives developed around innovative ideas on how to teach introductory programming using web technologies such as Code Avengers (http://www.codeavengers.com/) or CodeHS (http://codehs.com/). Web development is also present in undergraduate and graduate curriculum such as (ACM/IEEE, 2013), (ACM/AIS, 2010), (ACM/IEEE, 2008).

From the above cited works, the interest in the subject is evident. However, an analysis of the curricula reveals that it is difficult to cope with all the content and technologies within a single course, especially in high school or in introductory university courses. As proposed by (Connolly, 2011) and (Maiorana, 2013 (a)), the topics should be covered in three courses: basic markup languages and Cascading Style Sheet in the first course; client side programming in the second course; and server side programming in the third course, with usability and interface design issue spanning all three courses. Other research such as (Wang, 2009) points to the

need for subdividing the web programming curriculum into more courses, suggesting a two-course curriculum.

This work presents a curriculum suited for either a university course on web programming or for a third-level high school course, in accordance with the CSTA curriculum. The curriculum was field tested in a high school at the $13^{th}/13^{th}$ grade with students enrolled in a Computer Science specialization course. The curriculum, by leveraging on prerequisite database knowledge such as Entity Relationship model (E/R) and Data Manipulation language (SQL), focuses on server side web programming and offers the possibility, using simple examples and guided practical lab sessions, to present and practically apply, as well as verify, advanced database concepts such as stored procedures, transactions and security issues, both at the database and at the web application levels.

The paper is organized as follows: section 2 briefly reviews the literature on teaching web development; section 3 presents the context of the teaching experience, the pedagogical model, the contents and how the course was taught; section 4 discusses a first course evaluation, highlighting positive and negative aspects; and section 5 presents the conclusions and highlights future work.

2 STATE OF THE ART IN TEACHING WEB PROGRAMMING

In the research carried out by (Grove, 2007), the authors review ten years of research articles on teaching web development in colleges and universities. The authors report that the diversity of protocols and programming languages required for web development presents a problem. Mastering all the knowledge required is deemed "unreasonable" while some practical level of understanding in all the technologies is considered necessary. This conclusion correlates with the necessity to divide, especially in high school, the curriculum into at least three courses. This is reinforced by (Park, 2013) where a taxonomy of student errors in HTML and CSS is analysed, reinforcing the pedagogical idea that even entry level languages, not considered as "computationally expensive" as many other languages, require developing computational skills such as checking syntax errors and debugging. The work makes use of an educational tool, openHTML, a web editor developed by the authors. Another e-learning system designed for improving web programming skills, is presented in (Elgamal, 2013), and offers an integrated environment for web programming.

Among recent case studies and teaching experiences on web programming at the university level, it is possible to recall the following notable research: (Liu, 2011) presenting challenges and tools used; (Laverty, 2011) where the authors highlight the challenge of an "efficient delivery of a dynamic web development, database-driven platform"; (Baatard, 2007) presenting a course using the PHP language and addressing security issues; (Noonan, 2007) (Wang, 2006) and (Olan, 2009) presenting a course focused on server side programming and database interaction; (Tao, 2010) reporting on the positive effect of encouraging students to use code developed by others; (Hollingsworth, 2010) where the Google Cloud App was used along with servlets and Java Server Pages (JSP); (Gousie, 2006) reporting on an interdisciplinary approach to teaching web programming, graphics and design in a course for non-majors; (Stepp, 2009) presenting a CS 1.5 web programming course, and suggesting that educators present form and PHP scripting as early as possible; and (Adams, 2007) presenting a course developed around a Web project. The use of frameworks, recently reviewed in (Chao, 2013), is not recommended in introductory web programming courses.

From this brief literature review it is possible to highlight the following points:

- the difficulty and challenges in selecting the topics and providing sufficient depth (Liu, 2011)
- the rich set of technologies used, the lack of their integration and their development at a fast pace
- the importance of delivering a dynamic, database-driven web development
- the necessity to span the content of the topics among several courses and the interdisciplinary aspect of the topic.

This work reports a detailed practical experience suited for a web programming module and focuses on how these activities offer the possibility to practically explore and experiment with advanced database concepts such as stored procedure, transaction and security in the realm of web programming. The use of these advanced database concepts in web programming represents a unique case compared to other courses described in literature. The application of the blending of the proposed pedagogical approach, although

representing a well-established practice, is presented in a particular context and within a particular discipline and will hopefully provide an initial insight in the described setting.

3 CONTEXT, PEDAGOGICAL MODEL AND CONTENT OF THE EDUCATIONAL EXPERIENCE

This work presents a case study dealing with an educational experience in teaching web programming in an Italian high school. The class comprised 14 students, 12 male and 2 female, between the ages of 18 and 20. The students in Italy typically start school at the age of 6, with a few exceptions who start at 5, and after 13 years of school, finish at 19. Among the students, there were two who had to repeat several years of school. The class was the final course of a computer science curriculum lasting five years. In the last three years the students attended two types of courses each year: the first type of course centers on programming and software development and the second type of course centers on hardware and networking. In the third year programming course, the students have to learn basic programming, problem solving and algorithm design in an imperative language such as C, as well as basic web page design using HTML and CSS. In the fourth year the students learn object- oriented programming along with event driven programming. The students extend the web design topic by writing JavaScript code on the client side. In the fifth year the students in the programming course deal with Database design and SQL language in the first term and web programming in the second term. This content is accompanied by the second course where students learn networking architecture, both hardware and software, client server programming and security issues, such as cryptography. The reported teaching experience deals with the second term of the software course focused on web programming technologies, where the students are required to master database design with E/R diagram and SQL languages both for Data Manipulation and for querying. The course in the fifth year was divided into 28 weeks, 12 of which were dedicated to the web programming curriculum. The final two weeks allowed for individual and group project finalization. Each week consisted of a two-hour session for theory, ground discussions and assessment, and three one-hour lab sessions divided into two hours plus a one hour meeting, for a total of three meetings per week.

The course schedule for the first term is reported in Table 1. For a deeper discussion of the database curriculum the reader can refer to (Giordano, 2013 (c)).

Table 1: Course schedule for the first term.

Unit	Argument	Hours
1	E/R, entity, primary key, attribute	2
	Projection and filtering using one table. Regular expressions	6
	User defined functions	6
	User interface for parametric queries	4
2	Associations 1-N, 1-1; association attribute, multiple associations, association with role.	6
	Join and set operations	8
3	Associations N-N.	6
	Sub-queries. Nested and correlated sub-queries	6
4	Ternary and n-ary associations	8
	Group by and nested queries	6
5	Associative entity, Normalization.	2
	Having clause	4
6	Recursive associations.	6
	Total	70

The bibliography for students comprised an Italian book together with selected parts from two leading database books (Atzeni, 1999) (Elmasri, 2011), as well as teacher notes. The adopted pedagogical model, in accordance with modern constructivist theory (Duffy, 1992), was student-centered, where the students in each lab session are required to design and implement a solution to small problems, presented as specific questions requesting to perform a particular task. The questions guide the students to a final lab session goal of producing an application solving a specific problem or applying a new technology. An example of such questions for the 7th session, as listed in Table 3, is given in Table 2. This approach has the advantage of guiding students in the complex task of mastering different technologies, programming languages, as well as in learning and understanding the deep and fundamental concepts of database and web programming.

According to the revised Bloom taxonomy (Anderson, 2005), the questions guide the students in climbing all the levels of the taxonomy: students using self-directed work along with the lab assignment questions are guided in the cognitive process of remembering, understanding and

applying. More gifted students are able to analyze and evaluate, while a guided inquiry approach is used during the lab session to allow the other students to analyze and evaluate their solutions. The creative aspect is saved for the end of the lab, in the form of deeper questions, self-directed exploration and experimentation and, most importantly, in the final project, where it is suggested that the students apply all the course knowledge in building a complete, self-chosen web application. The final project can also be done in self-selected groups of students.

Table 2: An example of a lab session assignment.

N	Questions
1	Create, using the SQL language a database called store.
2	Create, using the SQL language, the following logical model: Person (id, Lastname, Firstname, age, sex, BornDate, salary, StartTime, CivilStatus) Login (id, idP, username, password) where CivilStatus can only have a value chosen from (single, nubile, conjugated, separated)
3	Insert data in the Person and login table. Verify that the reference integrity constraint works.
4	Create a stored procedure to insert a person. Create a function to insert a person returning the ID of the last inserted person. Verify that the stored procedure and stored function work by calling them
5	Create a stored procedure to insert in the login table and verify it. Verify that the reference integrity constraint works.
6	Create an HTML page containing a form to insert the data of a person including username and password.
7	Create a PHP page to insert in the database the data relative to the person including the login information. Verify that the transaction mechanism works by forcing the insertion in the login table of a row with a non-existent value for the foreign key.
8	Write a stored procedure that, by using the transaction, inserts a row in the person and login table, thus inserting all the information pertinent to the person. Simplify the PHP code in order to perform a single call to this stored procedure. Compare the solutions developed in steps 7 and 8.

The lab sessions were designed in an incremental way around the management of the login process and user registration. It was decided, for didactic and security reasons, to divide the user data and the login data (username and password) into two tables allowing the students to present and practically apply important concepts such as stored procedure, transaction, security issues and security advantages in using stored procedure.

A modern object-oriented API such as the Portable Database Object (PDO) was used for database access. The login process allowed students to deal with the main aspects of web programming: HTML, forms, PHP, SQL, session and cookies to keep track of the data across the web pages; data validation took place both on the client side using JavaScript and regular expressions, and on the server side using PHP function and database stored procedures. The choice of the model allowed for the application and verification, with few lines of code, of sophisticated mechanisms such as transactions and stored procedure invoker and creator security models, as well as SQL commands for security management. Possible lab extensions to the basic login page could provide forgotten password functionalities, date and time of last login, count of unsuccessful logins on the basis of the username and a personalized welcome message, just to name a few.

Throughout the entire term an inverted classroom approach (Gannod, 2008) was used where the students were asked to pre-read the material, thereby leaving time for practical work during the lab sessions and discussion during the theory sessions. These discussions included solutions to exercises and case studies with critical comparisons of different approaches to solving the problems. The material was based on the textbook as well as the teacher's notes. Modern web 2.0 technologies (www.blogger.com/) (sites.google.com/) were used to deliver content, encouraging students to use wiki technologies to develop a personal e-portfolio where solutions to case studies and projects were posted to allow peers and instructors to leave formative feedback (Giordano, 2004), (Giordano, 2013 (a, b). The free Easy PHP (http://www.easyphp.org/) Windows Apache MYSQL PHP (WAMP) was used on a Windows platform. Table 3 reports the course schedule with the major lab assignment topics.

The lab sessions were very fruitful and allowed for student self-discovery. For example, during a lab session a student working on a machine with an early version of PHP verified that the foreign key constraint does not work with the MYSQL MyISAM engine. It was suggested that the students use the manual to find a way to change the engine and verify that the foreign key constraint is implemented in other engines such as the InnoDB. The side effect was to push students to improve technical manual reading capabilities, particularly for SQL syntax such as DML instructions (MySQL, 2013).

This approach was extended to the PHP manual (PHP, 2013), thus improving a technical skill required in working environments. Overall, the

Table 3: Course schedule and lab assignments.

N	Topic	Lab assignment	#h
1	Three tier architecture	Install and configure a WAMP framework	3
2	HTML review	Login form creation	3
3	Database connection and data insertion through stored procedure.	Stored procedures for insertion in one table and their use with PDO. Parameters. Insert into the person table.	6
4	Stored procedure for updating and deleting data.	Use of HTML forms and stored procedure to update and delete data. Manage the person and login table. Get and POST method.	6
5	Stored procedures for retrieving data with and without parameters.	Retrieving data using stored procedures and PDO and presenting them in HTML table. CSS review. List of registered users.	6
6	Transaction: begin, commit and rollback	PHP code for managing transaction. SQL code for managing transaction. Insertion in the Person and login table	8
7	Use of transaction to insert, update and modify data across multiple tables	Use of transaction to insert, modify data in multiple tables. Management of the person and login table	4
8	Cookies and sessions	Keep track of username and ID across multiple web pages	6
9	SQL command for security	Granting and revoking privileges to users. Stored procedure for granting and revoking privileges	4
10	Security model of stored procedures: definer and invoker model. Cryptography.	Use and comparison of different security models to automatically manage user creation and management. Cryptography functions in PHP.	6
11	Client side form validation using JavaScript and regular expressions	Validate the Person and login data at the client side	4
12	Server side data validation.	Using PHP and stored procedure to validate the person and login data at the server side	4
		Total	60

approach was useful in guiding the students in the complex process of designing web applications, allowing for an incremental project that pushed the students to develop limited but working projects in each lab session, refining the solution throughout the semester. This approach avoided teaching each topic individually but rather fostered step-by-step teaching, through incremental integration.

Initially, the students were reluctant to follow the approach due to the added level of abstraction introduced by the stored procedures and hence the higher cognitive effort required to master them. While proceeding with this approach, their opinions changed positively and the students appreciated the benefit of the method in terms of performance and security. Likewise, they appreciated the reduced burden in writing PHP code and the possibility to resolve the basic database management task with the same approach and by using a similar design procedure.

The teachers' evaluation of the experience, by comparison with previous classes with the same coursework and the same teachers, where more emphasis was given to physical database design, indexing and query performance (Maiorana (b), 2013) thus leaving less time for more extensive web programming, was positive, mainly due to the possibility of exposing the students to a richer set of technology with a tighter link to database technologies and concepts. Overall, 6 students were able to complete all the guided lab exercises independently, with little guidance in the first lab session, 2 students completed the lab session but required more support from the teachers, 6 students struggled due to the difficulty and deficiency both in concepts such as procedure and functions that were covered in the previous years, and to a deficiency in the mastering of some topics related to the first part of the course and finally 1 student was not able to complete the course and pass the year due to the numerous insufficient results in other disciplines. The final student grades in the discipline were: 4 students received an A, 3 students a B, 5 students a C, 1 student a D and 2 students an F.

4 COURSE EVALUATION AND DISCUSSION

The course was evaluated by using student and teacher feedback, obtained by informal discussion and formal colloquia inside the class meetings and by assessing and evaluating student work in each lab session. The feedback from students as well as teachers was positive.

Students were initially overwhelmed by the complication introduced by stored procedure and the

necessity to manage and discriminate different working environments such as the MYSQL database, the PHP scripting languages as well as HTML. When the initial difficulties were overcome, thanks to a series of lab sessions focusing on similar concepts and technologies (e.g., stored procedure for inserting, deleting, updating and retrieving data from a single table), the students particularly appreciated using stored procedure and advanced database concepts which allowed them to use less code, perform canonical tasks such as inserting, updating, deleting and retrieving data in a canonical and recurring way, changing only the SQL command to be used in the stored procedure as well as the number and type of parameters.

From our colleagues' point of view, the approach allowed us to present and master advanced database concepts. The approach allowed for the sharpening and augmentation of database knowledge required by all the students in the final nation-wide state examination. The step-by-step approach avoided students getting lost in the multiple environment scenario typical of Web programming, despite the coverage of advanced topics.

The approach allowed students to experiment and easily verify important topics, such as transactions, by verifying their effects. Examples of this verification approach include: inserting a violating reference key data as the second step in a transaction, thereby allowing students to verify the missing effect in the first step; looking for the data stored on the client side (the cookie) and on the server side (the session file), thus allowing students to distinguish between the two concepts and their effects; verifying the different foreign key mechanisms with different MYSQL engines, and consequently mastering this important topic.

The assessment of the lab sessions revealed, through analysis of the results and of the grades assigned, three levels of achievement: the more gifted students were the first to proceed alone and expand the proposed lab session assignments; a second group of students were able to follow the steps and work through a solution to each step with minimal guidance; finally, a third group of students required particular attention and support to work through the exercises, with some of them falling behind and not meeting the deadlines. This division was reflected in the final grade of the module, as noted before. Despite the increased complexity of the course topic, the results correlate with and are slightly better than past experiences.

Overall, compared to a previous course taught by the same teachers, a deeper mastering of the web programming material was observed. Whereas only two students from the previous course were able to apply and discuss stored procedures and transactions in their final project (Maiorana (b), 2013), up to eight were able to do so in this course. This was due to the different teaching approach used, one that did not use the step-by-step assignment method in the lab session and which also dedicated fewer hours to the web programming module in order to explain another important topic, i.e., indexing, which is now an elective course in the new CS curriculum (ACM/IEEE, 2013).

5 CONCLUSIONS

A web programming curriculum was proposed, based on the results of a teaching experience in the final year at an Italian high school. The pedagogical approach was based on lab sessions requiring students to design and implement a solution with a small sequence of steps leading to a working artifact. The lab sessions subsequently refined each other by presenting and applying new concepts or techniques and incrementally improving a project spanning all the lab sessions which centered on the management of the login information on a website. In solving the login problem, the students were able to cover all the major topics of web programming, linking them to advanced database concepts such as transaction, stored procedure and security models, topics not usually covered in a high school or undergraduate setting. At the same time the concepts, mastered by the majority of the students, allowed for an easy framing of a solution pattern for the management of data in Web applications. A preliminary course assessment has also been presented.

The presented curriculum, according to Maiorana (2013 (a)) is suited for a third-level course dealing with server side programming. The course represents the concluding phase, following a tag-based language and Cascading Style Sheet in the first course and client side programming in the second course. The proposed approach presents the advantages of a strong relationship with database concepts, allowing at the same time to present initial issues related to web interfaces and usability.

The proposed curriculum and pedagogical approach can be used in both high school and university courses, allowing the students to present and practically apply all the concepts suited for the third-level curriculum in the proposed three-level course architecture. Regarding further research, the following is planned: a deeper qualitative and

quantitative course assessment compared to a previous class with the same coursework and the expansion of the curriculum by using frameworks, such as jQuery Mobile (http://jquerymobile.com/), to develop mobile web applications (Costanzo, 2013).

REFERENCES

Adams, D. R., 2007. Integration early: a new approach to teaching web application development. *Journal of Computing Sciences in Colleges*, Vol. 23, No. 1, pp. 97-104.

Anderson, L. W., Krathwohl, D. R., Bloom, B. S., 2005. *A taxonomy for learning, teaching, and assessing*. Longman.

Association for Computing Machinery (ACM), Association for Information Systems (AIS)**.** IS 2010 Curriculum Guidelines for Undergraduate Degree Programs in Information Systems. Retrieved on August 2013 from http://www.acm.org/education/curricula/IS%202010%20ACM%20final.pdf.

Association for Computing Machinery (ACM), IEEE Computer Society. Information Technology 2008. Curriculum Guidelines for Undergraduate Degree Programs in Information Technology. Retrieved on August 2013 from http://www.acm.org/education/curricula/IT2008%20Curriculum.pdf.

Astrachan, O., Briggs, A., 2012. The CS principles project, *ACM Inroads*, Vol 3, No. 2, pp. 38-42.

Atzeni, P., Ceri, S., Paraboschi, S., Torlone, R., 1999. *Database Systems: concepts, languages and architectures*. McGraw-Hill, 1999.

Baatard, G., 2007. Teaching PHP with security in mind. In Proceeding of the 5th *Australian Information Security Management Conference*, pp. 21-27.

Bureau of labor statistics. Occupational outlook handbook. Retrieved September 2013 from http://www.bls.gov/ooh/computer-and-information-technology/ information-security-analysts-web-developers-and-computer-network-architects.htm.

Chao, J., Davey, B., 2013. Navigating the Framework Jungle for Teaching Web Application Development. *Issues in Informing Science and Information Technology*, 10, pp. 95-109.

Connolly, R., 2011. Awakening Rip Van Winkle: modernizing the Computer Science Web Curriculum. In Proceedings of the 1*6th annual joint conference on Innovation and technology in computer science education (ITICSE, 2011)*, pp. 18-22, ACM, 2011.

Connolly, R. W., Miller, C. S., 2013. Teaching Web Development. Special Issue of the *ACM Journal Transactions on Computing Education*. Retrieved on August 2013 from http://toce.acm.org/attachments/OCE-SE-TW.pdf.

Costanzo, A., Faro, A., Giordano, D., 2013. WI-CITY: living, deciding and planning using mobiles in Intelligent Cities. In *3rd International Conference on Pervasive and Embedded Computing and Communication Systems, PECCS, Barcelona, INSTICC*.

CSTA Standards Task Force. 2011. K–12 Computer Science Standards Revised 2011. The College Board, Computer Science: Principles Big Ideas and Key Concepts Learning Objectives and Evidence Statements.

Duffy, T. M., Jonassen, D. H., (Eds.), 1992. *Constructivism and the technology of instruction: A conversation*. Psychology Press.

Elgamal, A. F., Abas, H. A., & Baladoh, E. S., 2013. An interactive e-learning system for improving web programming skills. *Education and Information Technologies*, *18*(1), pp. 29-46.

Elmasri, R. & Navathe, S., 2011. *Fundamentals of Database Systems* (6th Edition). Addison Wesley.

Giordano, D., 2004. Shared values as anchors of a learning community: A case study in information systems design. *Journal of Educational Media*, 29, 3, pp. 213-227.

Giordano, D., Maiorana, F., 2013 (a). An Interdisciplinary Project in Sustainable Development Based on Modern Visual Programming Environments and Web 2.0 Technologies. Proceedings of the *3rd Interdisciplinary Engineering Design Education Conference, IEDEC* 2013, pp. 163-167.

Giordano, D., Maiorana, F., 2013 (b). Giordano, D., Maiorana, F., 2013. Object Oriented Design through Game Development in XNA. Proceedings of the *3rd Interdisciplinary Engineering Design Education Conference, IEDEC* 2013, pp. 51-55.

Giordano, D., Maiorana, F., 2013 (c). Teaching database: a pedagogical and curriculum perspective. Proceedings of the *International Conference on Information and Communication Technology for Education, (ICTE)*, in press, 2013.

Gousie, M. B., 2006. A robust web programming and graphics course for non-majors. *ACM SIGCSE Bulletin*, Vol. *38*, No. 1, pp. 72-76.

Grove, R. F., 2007. Trends in teaching web-based development - A survey of pedagogy in web development courses. Proc. 3rd *International Conference on Web Information Systems and Technologies, Webist*, Barcelona, Spain.

Gundotra, V., 2009. Google I/O 2009 Keynote Day 1. http://www.youtube.com/watch?v=S5aGZIvk&feature=channel.

Gannod, G. C., Burge, J. E., Helmick, M. T., 2008. Using the inverted classroom to teach software engineering. In Proceedings of the *30th international conference on Software engineering*, pp. 777-786, ACM.

Hollingsworth, J., Powell, D. J., 2010. Teaching web programming using the Google Cloud. In Proceedings of the *48th Annual Southeast Regional Conference*, pp. 76-80, ACM.

Laverty, J. P., 2011. Implementing a dynamic database-driven course using LAMP. *Information System Education Journal (ISEDJ)*, Vol. 9, No. 7, pp. 33-40.

Liu, Y., Phelps, G., 2011. Challenges and professional tools used when teaching web programming. J*ournal*

of Computing Sciences in Colleges, Vol. *26*, No. 5, pp. 116-121.

Maiorana, F., 2013 (a). Teaching web programming: critical considerations on pedagogical approach and curriculum content, Proc. of International Conference on Sport Science and Computer Science.

Maiorana, F., Giordano, D., 2013 (b). A constructivist approach to teaching index selection strategies and database design. Proceedings of the *International Conference on Information and Communication Technology for Education, (ICTE)*, in press, 2013.

MySQL 5.6 Reference Manual. Retrieved August 2013 from http://downloads.mysql.com/docs/refman-5.6-en.a4.pdf.

National Science Foundation (US) National Science Board. 2012. Science and Engineering Indicators.

New Zealand Ministry of Education, 2010. Digital technology guideline, dtg.tki.org.nz.

Noonan, R. E., 2007. A course in web programming. J*ournal of Computing Sciences in Colleges*, Vol. 22, No. 3, pp. 23-28.

Olan, M., 2009. Web applications: a test bed for advanced topics. *Journal of Computing Sciences in Colleges*, Vol. 24, No. 3, pp. 72-80.

Park, T. H., Saxena, A., Jagannath, S., Wiedenbeck, S., Forte, A., 2013. Towards a taxonomy of errors in HTML and CSS. In *Proceedings of the ninth annual international ACM conference on International computing education research* (pp. 75-82). ACM.

PHP manual. Retrieved August 2013 from http://us3.php.net/get/php_manual_en.html.gz/from/a/mirror.

Stepp, M., Miller, J., Kirst, V., 2009. A CS 1.5 introduction to web programming. In *ACM SIGCSE Bulletin* Vol. 41, No. 1, pp. 121-125, ACM.

Tao, Y., Plouff, C., 2010. Work in progress: An opportunistic approach for learning web application development. In *Frontiers in Education Conference (FIE)*, 2010 IEEE (pp. S2D-1), IEEE.

The Computing at School Working Group. Computer Science: A curriculum for schools (2012). Retrieved August 2013 from http://www.computingatschool.org.uk/data/uploads/ComputingCurric.pdf.

The Joint Task Force on Computing Curricula Association for Computing Machinery IEEE-Computer Society. Computer Science Curricula 2013. Ironman Draft. Retrieved Agust 2013 from http://ai.stanford.edu/users/zsahami/CS2013/ironman-draft/cs2013-ironman-v1.0.pdf.

Wang, X., 2006. A practical way to teach web programming in computer science. *Journal of Computing Sciences in Colleges*, Vol. 22, No. 1, pp. 211-220.

Wang, Y. D., Zahadat, N., 2009. Teaching web development in the Web 2.0 Era. In Proceedings of the *10th ACM SIGITE conference on Information technology education*, pp. 80-86.

The Elevate Framework for Assessment and Certification Design for Vocational Training in European ICT SMEs

Iraklis Paraskakis[1] and Thanos Hatziapostolou[2]
[1]South East European Research Center (SEERC), International Faculty of the University of Sheffield, City College, 24 Proxenou Koromila, 54622, Thessaloniki, Greece
[2]Department of Computer Science, International Faculuty of the University of Sheffield, City College, Leontos Sofou, 54632 Thessaloniki, Greece
iparaskakis@seerc.org, a.hatziapostolou@city.academic.gr

Keywords: Vocational e-Training, Competence-based Education, e-Assessment, Web2.0.

Abstract: This paper discusses vocational training and e-Training within the context of IT SMEs, focusing on the process of assessment and certification. More specifically, current trends in the assessment and certification of IT skills are discussed, as revealed in the study of related work. In addition, issues in traditional and e-Assessment are presented, and a new approach: "Assessment 2.0", which exploits the characteristics of contemporary trainees, is proposed. Finally, after examining relevant research projects, and current European certification programs, the ELEVATE project approach to certification and assessment is introduced with an example. More precisely, the proposed methodology for defining competence-based learning is put forward.

1 INTRODUCTION

The continuing Information and Communication Technology (ICT)-driven evolution of products and processes, coupled with the need for a low-carbon economy as well as population ageing, will mean that jobs and social structures will change: education and training, including vocational education and training (VET), must adapt accordingly (COM 296, 2010). The Europe 2020 Strategy puts a strong emphasis on education and training to promote "smart, sustainable and inclusive growth" and to reinforce the attractiveness of VET.

Actions to improve VET help to provide the skills, knowledge and competences needed in the labor market. As such, they are an essential part of the EU's 'Education and Training 2020' work programme. Also, encouraging learners to take part in VET in different countries is also a priority of EU actions, providing individuals with increased opportunities and experiences, and enhancing efficiency and innovation.

Assessments are the foundation of effective instructional practices and return-on-investment studies, since in research such as that by Glahn (2008), it is revealed that assessment does not only allow the expression but also the comparison of knowledge and competences among groups of learners. Moreover, the power and consequences of assessment have become exponentially more important with the advent of content management systems (CMS) and learning management systems (LMS) which foster communication and dissemination.

At its most basic level, assessment is the process of generating evidence of trainee learning and then making a judgment about that evidence (Elliott, 2008). Current assessment practice provides evidence in the form of examination scripts, essays and other artefacts. Furthermore, data from assessments helps drive the development of solid content and advances instructional practices.

'Assessment 1.0' can be thought of as assessment practice from the beginning of the 20th century until today. Throughout this period, assessment exhibited the following characteristics:

- paper-based,
- classroom-based,
- formalised (in terms of organisation and administration),
- synchronised (in terms of time and place), and
- controlled (in terms of contents and marking).

According to the Europe 2020 Strategy, there appears to be a need for greater flexibility regarding

how learning outcomes and competences are acquired, how they are assessed and how they lead to qualifications. In formal educational settings, for example, this is achieved through specially trained staff and the application of fine grained assessment methods. Methods which are not limited to formally strict testing, but can include observing learners on their learning course, stimulating group work, apprenticing, analyzing a learner's contributions in discussions and problem solving approaches.

A more up-to-date form of assessment has emerged in the last decade, which involves the use of computers in the assessment process (Elliott, 2008). 'E-Assessment' also embraces 'e-testing' (a form of on-screen testing of knowledge) and 'e-portfolios' (a digital repository of assessment evidence normally used to assess practical skills). This approach to assessment runs tangent to the application of Software as a Service (SaaS).

SaaS is going to be a big part of the way we work and learn in the future (Marks, 2008). Some of the advantages of SaaS include: a) lower maintenance and functionality costs, b) no need for software licenses or hardware upgrades, c) increased mobility since documents are accessible anywhere, through the Internet, and d) documents are safely stored remotely.

According to Walker et al. (2004), there are six key components for the successful delivery of e-Assessment: central support, quality software, quality hardware, clear policies and procedures, integration within the learning system and staff education. Therefore, considering the advantages of utilizing SaaS, the burdens of implementing e-Assessment for an SME could be limited to just integration with an existing learning system and staff education.

Firms generally praise online training (Strother, 2002) as a cost-effective, convenient, and productive way to deliver corporate education. Results of studies, such as by Hamburg et al. (2008) however, show that less than 25% of SME (Small and Medium Enterprises) staff participates in vocational training courses and less than 60% of employers provide any type of training for their staff. This is mainly due to the fact that many SMEs have neither enough knowledge, or resources to develop and implement sustainable training strategies based on new media and knowledge processes for their own organization. Issues such as this will be analyzed in a following section, followed by appropriate proposed solutions.

One of the key qualities of formal education is to make learning processes accessible for communication in such a way that forms of assessment results (i.e., certificates) can serve as proof of the acquired knowledge and competences. This paper focuses on the certification of IT skills in SMEs and presents the rationale behind the certification approach selected for the ELEVATE project members.

In the rest of the paper, we examine related research on assessment in vocational training, focusing on the assessment of IT skills in SMEs. The paper describes current practices, as well as issues pertaining to traditional and electronic assessment. Next, we examine an update to current assessment practices, which utilizes Internet technologies and, more specifically, Web2.0. This is referred to as Assessment 2.0. Following this, the paper presents the assessment and certification approach of the ELEVATE project and describes an example of its application. Finally, the last section discusses conclusions and future work.

2 ISSUES ON ASSESSMENT AND CERTIFICATION IN SMES

Information technology proficiency among citizens is a key factor in the dynamics of Information Society development and further economic growth. In the following sections we examine specific issues which hinder the processes of assessment and certification in SMEs, and discuss related projects and initiatives.

2.1 Traditional and e-Assessment Issues

In traditional types of assessment (Assessment1.0), the formats of the examinations range from 45-120 minute restricted response (i.e., multiple-choice, identification, one best correct answer, 2/3/4 correct answers), to restricted-response adaptive modes (that stop the testing process at the point in the adaptive curve at which a passing score could be predicted at a 95 percent confidence level), constructed response (e.g., drag and drop), and "essays", simulations, scenarios and case studies involving performance benchmarks.

Traditional assessment has a bureaucratic nature, which is expensive to run and doesn't scale well (Elliott, 2008). It's also inflexible and organised around annual examination 'diets'. Moreover, some educationalists claim that the current assessment system encourages surface learning and "teaching to the test". Instead of instilling genuine problem solving skills, it fosters memorisation.

The traditional assessment approach concentrates

mainly on the testing of basic skills, supposedly acquired through mainly drill and practice experiences. Such an assessment system is often referred to as a test culture (Sluijsmans, 2002). On the contrary, assessment that is performance-oriented, such as the ELEVATE project's competence-based approach, aims to measure not only the correctness of a response, but also the thought processes involved in arriving at the response.

Employers in SMEs complain that, in spite of rising achievement (DfES, 2007), young people are not gaining the skills that are needed in the modern workplace – skills such as problem solving, collaboration, innovation and creativity. Vocational education and training must equip young learners with skills directly relevant to evolving labor markets, such as e-skills, and highly developed key competences; such as digital and media literacy to achieve digital competence (COM 296, 2010).

In addition, trainers complain about the rising burden of time spent carrying-out and marking assessments, which reduces the time available for "real teaching". These criticisms are not confined to paper based assessment. E-testing has been criticised for crudely imitating traditional assessment. These criticisms of e-Assessment mirror the criticisms of virtual learning environments (VLEs) – that they simply seek to mimic traditional classroom practice.

Both paper-based and computer-based assessments are perceived by trainees as something external to them; something over which they have no control; something that is 'done' to them. And the assessment instrument itself is considered contrived, just a hurdle to be jumped, not part of their learning. Or, worse, it is perceived as the sole purpose of their learning, with all their efforts going into passing the test rather than the acquisition of new knowledge and skills.

Assessment 1.0 is also intensely individualistic. Assessment activities are done alone, competition is encouraged, and any form of collaboration is prohibited. Furthermore, the use of e-Assessment systems might hold back progress in assessment by constraining practice to traditional (paper based) assessment and the limited form of computer-based assessment made possible by these systems.

2.2 Issues on the Certification of IT Skills

The most basic principle of IT industry certification is that content counts. Given the global reach of IT certification, some examinations are available in languages other than English, but that depends on the subject matter.

Certification tests in the IT world are constantly being retired and replaced to meet the current state of vendor products and industry knowledge (Adelman, 2000). Testing firms do not award the formal certifications. Rather, they report examination results to vendors and industry associations that issue the documents of certification, and both vendors and industry associations may have requirements for certification beyond those of examination.

Increasingly, too, we find "cross-vendor recognized" examinations, a development that underscores the rationalizing trends in the industry, along with the competition for trained labour. Microsoft, for example, waives its networking examination requirement for those who are already certified by Novell, Banyan or Sun as network engineers, specialists or administrators. These cross-vendor recognized examinations are a prelude to the adoption of industry-wide certification standards and accreditation.

2.2.1 Related Projects

During research, the ELEVATE project discovered some highly relevant projects, which utilize competence-based training. In the following paragraphs, these projects are examined with respect to their assessment and certification approaches:

- PROLIX – Although studying the PROLIX research findings enabled the ELEVATE project to gain some important insight into competence-based training, there doesn't seem to be enough information on how these competences are assessed. Certainly, specific types of assessment are defined, but there isn't enough available information, to allow the ELEVATE partners to analyze, evaluate, adapt and perhaps integrate specific methodologies.
- TenCompetence – The TenCompetence project has yielded some interesting results, especially in the form of pedagogical tools, aiding Learning Design based on the IMS-LD standard. Yet, TenCompetence does not seem to focus on the certification of competences. It seems the validation of competences is realized outside of the system, by specific organizations. The user simply submits this certification as evidence of acquired competencies. In other words, there is no integration of the certification or assessment process within the TenCompetence system. Therefore, the TenCompetence system can be mostly utilized as a learning design tool for the

trainer and an e-portfolio mechanism for the trainee.

- The Leonardo da Vinci Programme – As part of the Lifelong Learning Programme, it funds a wide range of actions in vocational training, ranging from opportunities for individuals to improve their work-related skills through placements abroad, to co-operation projects between training organisations in different countries. With regards to certification, mobility certificates are given to organisations that have shown a particular quality in carrying out Leonardo da Vinci mobility projects (knowledge, experience and resources) and have developed an internationalisation strategy. Therefore, this Programme does not support the certification of individuals, as required by the ELEVATE SME certification requirements. However, there are a number of initiatives under development to enhance the transparency, recognition and quality of competences and qualifications, facilitating the mobility of learners and workers. These include the European Qualifications Framework (EQF), Europass, the European Credit System for VET (ECVET), and the European Quality Assurance Reference Framework for VET (EQAVET). From the initiatives mentioned above, only EQF and Europass are relevant and will be examined in a following section.

2.2.2 European Certification Programs

During research, four European initiatives relevant to certification have been identified: EITCI, ECQA, EQF, and Europass. These initiatives are presented in more detail below.

- Europass is an EU initiative to increase transparency of qualification and mobility of citizens in Europe. It aims to be a Life Long Learning Portfolio of documents containing the descriptions of all learning achievements, official qualifications, work results, skills and competencies, acquired over time, along with the related documentation. Therefore, this is not an approach to certification, but a form of e-portfolio. However, by following the specific standards and templates set by Europass, the ELEVATE user profile should be able to be adapted and exported.
- The European Information Technologies Certification Institute (EITCI, http://www.eitci.org/) has been established as a not-for-profit European, non-governmental organization, dedicated to counteracting digital exclusion in society by undertaking research and development of IT certification methodologies and standards. The EITCI Institute currently runs two IT certification programs: a) the European IT Certification Course EITCC program, and b) the European IT Certification Academy EITCA program. Both certification programs are accessible over the Internet and are targeted at individuals and institutions, enabling formal documentation of information technology competencies. Due to a specially designed fully on-line certification procedure preserving high quality of the certificates, there is no need to undertake EITCC and EITCA examinations at a physical location. This is an important factor in overcoming individual barriers and ensuring accessibility of IT certification services to everyone. Furthermore, within the framework of the EITC Program, the Institute is currently supporting the following IT Professional Certification Paths: a) *EITC-I: Internet Technologies* - This certification path is recognized globally as a formal proof of applicable knowledge, qualifications and expertise in the Internet and eCommerce technologies domain, b) *EITC-S: Security Technologies* - This certification path is recognized globally as a formal proof of applicable knowledge, qualifications and expertise in the IT security technologies domain, c) *EITC-M: Information Management* - This certification path is recognized globally as a formal proof of applicable knowledge, qualifications and expertise in the Information Technologies management domain, and d) *EITC-D: Software Development* - This certification path is recognized globally as a formal proof of applicable knowledge, qualifications and expertise in the areas of modern software development and engineering. Although EITC covers a number of certification paths, these are too general in order to be applicable to specific SME needs. In other words, SME certification types should not necessarily conform to a European standard but should be able to adapt and merge with it. This can be achieve through ECQA which is presented next.
- The ECQA (European Certification & Qualification Association, http://www.ecqa.org/) is the result of a number of EU supported initiatives in the last ten years, where in the European Union Life Long Learning Program

different educational developments decided to follow a joint process for the certification of persons in the industry. Through the ECQA it becomes possible to attend courses for a specific profession in one country and perform a Europe-wide agreed test at the end of the course. If an organisation or consortium wants to propose a new European profession for ECQA certification it must apply the guidelines for ECQA certified job roles and submit a job role proposal to the ECQA board members. Criteria for Certification of EU Job Roles include: a) the job role descriptions must comply with the European standards for skills descriptions. Each skill unit team describes the skill unit, learning elements and performance criteria following the standards for skills descriptions, b) a pool of test questions must be provided using a specific standard description format (so that the test questions can be entered into the exam portals), c) the pool of test questions must contain at least 5 different test questions per performance criterion so that a test per participant can be generated generically, d) to be accepted as a profession with a cross-regional and cross-national European scope the job role must additionally satisfy European dimension criteria, e) to be accepted as a profession with impact on a European level there must be a partnership agreement available (examples are consortium agreements, exploitation agreements, qualification boards) which assures that the profession and its skills card will be maintained for a number of years (minimum is 3), and f) to establish an independent and computer controlled test and certification centre all professions must use the European skills portals to administer a test pool and run all exams through it (with an independent objective board controlling the test servers and procedures).

The ELEVATE certification approach is based on the ECQA. The ECQA corresponds to the Bologna principles and European Qualification Framework (EQF). The EQF acts as a translation device to make national qualifications more readable across Europe, promoting workers' and learners' mobility between countries and facilitating their lifelong learning.

However, in order for the ELEVATE system to be in accordance with the ECQA specifications, the assessment and certification process should be carefully designed to meet specific criteria, as is described in the next section.

3 THE ELEVATE PROJECT APPROACH

Based on the bibliographical research and analytical study presented in the previous sections, we deduced the fundamental characteristics of the envisaged certification and assessment methodology of the ELEVATE educational system.

3.1 ELEVATE Project Assessment Approach

Educational innovations, such as problem and competence based education, are more likely to succeed if they include new forms of assessment, whereby assessment and learning are strongly interconnected in the course materials. Trainers in SMEs, however, usually lack the expert knowledge of pedagogists or instructional designers, required in order to facilitate this process.

For this reason the ELEVATE system incorporates a four step methodology, which allows the design of courses in which instruction and assessments are completely aligned:

- Define the purpose of the performance assessment.
- Choose the assessment task - Issues that must be taken into account are time constraints, availability of resources, and how much data is necessary in order to make an informed decision about the quality of a trainee's performance. The literature distinguishes between two types of performance based assessment activities that can be implemented: informal and formal.
- Define performance criteria.
- Create assessment forms - In these forms, trainers determine at what level of proficiency a trainee is able to perform a task or display knowledge of a concept.

The steps stated above define the necessary components for complying with the ECQA requirements., The approach could be further extended by including informal methods of assessment, used as evaluation factors to appointed performance criteria. Such informal methods of evaluation could utilize Web2.0 tools such as blogs, participation in the creation or editing of a wiki article, searching for external sources relevant to the studied topic, uploading documents/videos of the training experience or participating in online social activities such as forums, chat rooms, or social network groups.

Assessment2.0 is an approach proposed by Elliott

(2008), as an update to traditional assessment, best suited for the characteristics of contemporary learners. Trainees are already using Web2.0 services as part of their everyday lives. As a result, education and training approaches are becoming disconnected. The classroom is a drab, technology-free zone that bears little relation to the increasingly technological reality of people's lives outside of the classroom.

Furthermore, an Assessment 2.0 approach allows trainees to take more control of their own learning and to become more reflective. This was also identified as the vision of the JISC (Joint Information Systems Committee) for 2014, which funded an e-Assessment Roadmap, which reviewed current policies and practice relating to e-Assessment across the UK (Whitelock, 2007).

Unfortunately a wide range of examples that have used Web2.0 tools are not yet available in the literature. In the meanwhile, however, the way advances with Web and Web 2.0 tools are addressing the new assessment agenda have been evaluated. Four examples of summative assessment presented by Whitelock (2010), illustrate effective methods of web technologies supporting innovative practices that display a number of Elliott's key characteristics. Tried and tested pedagogical strategies have been enhanced in many of the cases examined and illustrate significant learning gains with the introduction of these technologies.

Assessment 2.0 poses challenges for trainers – who are often the epitome of the digital immigrant. Not only might they lack the IT skills needed to understand Web 2.0 services but they may lack the knowledge and experience required to appraise trainees' work produced using these tools. Furthermore, trainers also lack the rubrics required to assess Web 2.0 skills.

For the reasons stated above, the utilization of Assessment2.0 in the ELEVATE project approach is regarded within the context of informal assessment. In other words, trainees are not evaluated based on their participation in social network activities or on their skills with Web2.0 tools. For ELEVATE, Assessment2.0 is regarded as an optional, supplementary activity to regular training.

3.2 ELEVATE Project Certification Approach

The pedagogic approach to the vocational training adopted by the ELEVATE project is a competency-based development strategy. In order to achieve an efficient implementation of this approach, the ELEVATE project introduces the Competency Graph as the domain knowledge model. Since the competence graph is developed by the domain expert through a system component, the interface should support the design and implementation of all the necessary features of the suggested approach for assessment and certification.

As mentioned before, the ELEVATE certification approach is based on the ECQA. The ECQA corresponds to the Bologna principles and European Qualification Framework (EQF). Therefore, the main goal for the certification component of the system is to be compatible with the ECQA certification template and specifications.

In order for the system to achieve this, some features need to be designed and implemented:

- The domain expert must be able to define certifications for specific job roles or professions.
- For each job role, the domain expert must be able to designate the specific competencies which are considered requirements of certification.
- For each competency, the domain expert must be able to define multiple performance criteria (i.e., skills exhibited by the trainee which serve as evidence of competency acquirement).
- For each performance criterion, the domain expert must be able to define at least five (5) evaluation questions based on the principles set by the ECQA template, as well as a factor representing the percentage of success which qualifies for a pass grade.

In addition, the ELEVATE user profile should follow the specific standards and templates set by Europass in order to allow it to be adapted and exported for the specific e-portfolio system. This would necessitate the definition of these fields in the user profile: personal information (name, birth date, nationality, contact etc.), desired employment (this could be the suggested job role defined for the ECQA certification), work experience (this is arbitrarily set by the user), education and training (this is arbitrarily set by the user), personal skills and competences (this can include information set arbitrarily by the user and acquired competences based on the ELEVATE system and categorised into team work, mediating skills, intercultural skills, computer skills etc.).

3.3 Designing the ELEVATE Project Assessment and Certification Approach

CAS is a software company based in Germany and

one of the three SME participants of the ELEVATE project. CAS develops Customer Relationship Management (CRM) solutions and provides face to face training of its developed genesisWorld software to customers and partners.

Based on the CAS genesisWorld Marketing Module e-Training analysis, which resulted in the competency graph depicted below (Figure 1), we will provide an example of the suggested ELEVATE approach to assessment design and certification implementation. The design of the training and evaluation approach will follow a specific step-by-step methodology, conforming to the templates and principles set by the ECQA association. The process, realized by the domain expert, is as follows:

1. Define the job role (i.e., skill unit)
2. Define the participating elements (i.e., competencies)
3. Define the performance criteria (i.e., learning objects)
4. Define the test questions (i.e., formal assessment)
5. Define the informal assessment method

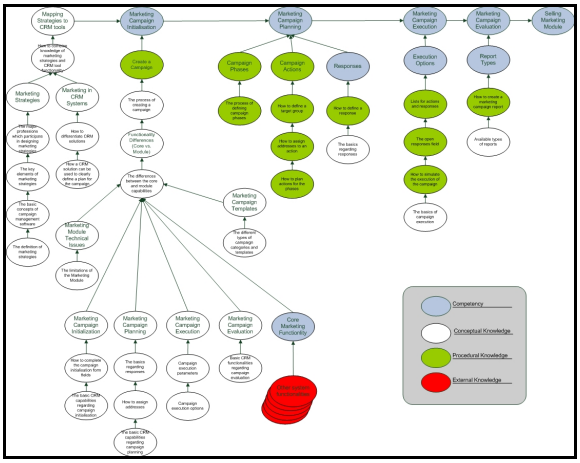


Figure 1: Competency Graph for Marketing Module Training.

3.3.1 Define the Job Role (I.E., Skill Unit, Certificate)

By adopting the ECQA rules and complying with set specifications, the ELEVATE project could propose a job role, in coordination with the participating SMEs, defined based on European standards for skills descriptions. Example job roles could include: CRM Sales-person, CRM Developer, CRM Administrator etc. In essence, the job role defined by the trainer here corresponds to the certification awarded after successful completion of the course contents.

Therefore, the first task of the trainer / domain expert is to define the job role, which will be taught to the trainees. The properties of the job role that must be defined include:

Skill Unit MM-SP: Marketing Module Sales Person – genesisWorld Sales Marketing Module

Skill Unit Acronym: MM-SP

Skill Unit Title: genesisWorld Sales Marketing Module

Skill Unit Description: This unit consists of five (5) elements: Marketing Campaign Initialisation, Marketing Campaign Planning, Marketing Campaign Execution, and Marketing Campaign Evaluation.

3.3.2 Define the Participating Elements (I.E., Competences)

Next, the tutor / domain expert has to define the required competences, which must be acquired by the trainee in order for him or her to be eligible for a specific job role or certification.

The CAS genesisWorld Marketing Module e-Training analysis resulted in the competency graph depicted in Figure 1. The goal Competency (i.e., the most general concept the trainee will learn) is the competency on the far right (i.e., Selling Marketing Module).

A competency consists of learning objects which can refer to either procedural or conceptual knowledge. When a competency only contains procedural learning objects it is coloured green and is referred to as procedural knowledge. On the other hand, if a competency only contains conceptual learning objects it is painted white and is referred to as conceptual knowledge. A competency which contains both types of learning objects is coloured light purple.

Finally, knowledge which is essential for acquiring the competency but which is not included in this part of the training is referred to as external and is coloured red.

3.3.3 Define the Performance Criteria (I.E., Learning Objects)

The Learning Objects, as used by the ELEVATE project are distinguished into categories based on the offered degree of interaction:

- **Passive:** The Learning Objects are presented to the trainee in the form of a lecture, without any kind of interaction between the trainee and the system (e.g., text manuals, video tutorials etc.).
- **Interactive:** These Learning Objects are

constructed for the purpose of interaction of the trainee with the training system, usually in order to diagnose the knowledge state of the trainee.

Passive learning objects can contain inherent degrees of scaffolding support for the trainee. These refer to the tactic of the training approach, and can be: a) implicit, and b) explicit.

The **explicit** approach (e.g., definition, example) is the simplest and most direct approach. In other words, no cognitive effort or reasoning capability is required on the part of the trainee, other than to memorise the information presented.

On the other hand, an **implicit** approach (e.g., analogy, example, discovery learning) to learning requires some form of deductive reasoning on the part of the trainee. For example, in the implicit approach of discovery learning a chapter in a book is given to the trainee and he/she has to mine the chapter for the necessary information.

Therefore, in this step the tutor defines the learning objects for each of the required competencies. Properties of the learning object which must be clearly outlined include: learning goal, id, type, source and presentation method. For example:
Learning Goal: Trainee learns the key elements of marketing strategies.
Identifier: MM-SP- MCI- MS- LO3a
Type: Passive | Implicit | Discovery Learning
Source: Book: "Customer Relationship Management" – pages 418-20
Presentation Method: Text-based document

3.3.4 Define Formal and Informal Assessment

After training for each competency, a small evaluation is carried out based on the learning objects presented and the knowledge acquired. Two types of evaluation can be defined: a) formal, and b) informal. **Formal** evaluation is based on traditional methods of assessment such as multiple choice questions, true/false statements, open ended questions etc. If the trainee fails at formal evaluation, a different training approach for the same learning object is presented.

On the other hand, **informal** assessment is based on contemporary tools of communication and collaboration such as wikis, social networks, forums, and document/video repositories etc. For example:

- Search the internet for examples of real life marketing strategies and post links of them on your personal blog.
- Define a detailed example of a fictional marketing strategy and attach it to the Examples section of the relevant wiki.

The completion of informal assessment activities is optional for the trainees, since their performance is not strictly evaluated. However, such activities enable trainees to participate in relevant communities and contribute to the development of additional training material for future generations of trainees.

3.4 Applying the ELEVATE Project Assessment and Certification Approach

The training begins with the trainee selecting a specific target competency he or she desires to acquire. Next, the ELEVATE system calculates an optimal learning path towards that goal (based on the defined competence graph of the domain knowledge) and presents the appropriate learning object to the trainee. If the trainee hasn't defined any training approach preferences, the default sequence is applied. The default sequence progresses from the more implicit approach to the most explicit definition. For example, a sequence can present: analogy, example, discovery learning and definition approaches in that order.

After studying a training approach, the trainee is presented with some feedback, with the system requiring a selection between two alternatives. The trainee can choose to examine the same learning object again, albeit with a different, more explicit, training approach. Alternatively, if the trainee feels confident enough, he or she can proceed to the assessment stage.

If the trainee selects to be evaluated, a set of formal assessment questions are presented. In order for the trainee to acquire the studied competence, he/she must answer correctly at least 75% of the questions. If the trainee is successful at the assessment, the system presents some optional informal assessment activities he/she can participate in. The trainee can then proceed with the next required competence along the predefined learning path.

If the trainee fails at evaluation, he/she is presented with the same learning object, through a different training approach. The selection of training approach is based on the default sequence or user preferences and on the questions which were answered incorrectly. If the trainee has studied all the learning object training approaches and still fails at evaluation a new methodology is attempted.

For each round of training-assessment-failure, the system presents the trainer with the default sequence of training approaches. The difference is that in each round, the most implicit training approach is ignored.

This method leads to a gradually increasing level of explicitness in the learning object training approaches. If the trainee continues to fail, even with the most explicit approach, the system restarts.

Finally, when the trainee completes the learning path, and acquires all the associated competences along the way, the system checks to see if any of those competences participate in the requirements of a certification. If the necessary competences have been attained, the system awards the certification. Alternatively, the system informs the trainee of his progress towards certification and suggests competences which should subsequently be acquired.

4 CONCLUSIONS & FUTURE WORK

This paper discussed vocational training and e-Training within the context of IT SMEs, focusing on the process of assessment and certification by examining relevant work, research and projects. Next we proposed a methodology, in order to aid SME in the design of the training, evaluation and certification approach, which conforms to the templates and principles, set by the ECQA association and Europass standard.

Currently, a prototype of the ELEVATE e-Training system is being tested and evaluated at the three participating SMEs. Based on the comments and suggestions of the evaluators we will ascertain the effectiveness and applicability of the proposed methodology and if necessary augment the e-Training system with desirable functionality.

REFERENCES

Adelman, C., 2000. *A parallel postsecondary universe: The certification system in information technology*. Education Publications Center, US Dept. of Education, ISBN: 0-16-050512-7.

COM 296., 2010. Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, (2010). Brussels, 9.6.2010 final, http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0296:FIN:EN:PDF.

DfES., 2007. *Johnson Welcomes Rising Achievement*. Retrieved 22 April 2007. http://www.dfes.gov.uk/pns/DisplayPN.cgi?pn_id=2006_0119.

Elliott, B., 2008. Assessment 2.0, *Proceedings of the Open Workshop of TenCompetence, Empowering Learners for Lifelong Competence Development: Pedagogical, Organisational and Technological Issues*, Madrid – España, April 10-11, 2008.

Eraut, M. (1997. *Developing Professional Knowledge and Competence*. The Falmer Press, London (1997).

Glahn, C., 2008. Supporting reflection in informal learning. *Presented at the ECTEL Doctoral Consortium '08*. September, 16, 2008, Maastricht, The Netherlands.

Hamburg, I., Engert, St., Anke, P., & Marin, M., 2008. Improving e-Learning 2.0-based training strategies of SMEs through communities of practice. Proceedings of the Seventh IASTED International Conference, Web-based Education, March 17-19, 2008 Innsbruck, Austria.

Holtham, C., Rich, M., 2008. Developing the architecture of a large-scale informal e-Learning network, *Proceedings of the Open Workshop of TenCompetence, Empowering Learners for Lifelong Competence Development: Pedagogical, Organisational and Technological Issues*, Madrid – España, April 10-11, 2008.

LPC., 1995. *Beroep in beweging. Beroepsprofiel leraar primair onderwijs* [Profession in action. Vocational training profile for the primary school teacher]. Utrecht: Forum Vitaal Leraarschap.

Marks, G., 2008. Beware the Hype for Software as a Service, http://www.businessweek.com/technology/content/jul2008/tc20080723_506811_page_2.htm.

Schön, D. A., 1983. *The Reflective Practitioner: How Professionals think in Action*. Maurice Temple Smith, London (1983).

Sluijsmans, D. M. A., 2002. *Student involvement in assessment. The training of peer assessment skills*. Unpublished doctoral dissertation, Open University of the Netherlands, The Netherlands, http://dspace.ou.nl/handle/1820/1034.

Strother, J., 2002. An Assessment of the Effectiveness of e-Learning in Corporate Training Programs, *International Review of Research in Open and Distance Learning*, ISSN: 1492-3831, Vol. 3, No. 1 (April, 2002).

Walker, D., Adamson, M., Parsons, R., 2004. Staff Education – Learning about Online Assessment, Online. In: Danson, M., ed., *Proceedings of 8th International CAA Conference*. Loughborough, University of Loughborough.

Whitelock, D., 2007. Computer Assisted Formative Assessment: Supporting Stuidents to become more Reflective Learners. *Proceedings of the 8th International Conference on Computer Based Learning in Science*, CBLIS 2007, 30 June – 6 July 2007. (eds.) Constantinou, C.P., Zacharia, Z.C. and Papaevripidou, M. pp.492-503 ISBN 978-9963-671-06-9.

Whitelock, D., 2010 Activating Assessment for Learning: are we on the way with Web 2.0? In M.J.W. Lee & C. McLoughlin (Eds.) Web 2.0 - Based-E-Learning: Applying Social Informatics for Tertiary Teaching. IGI Global.

Whitelock, D., Brasher, A., 2006. Developing a roadmap for e-Assessment: Which way now? Paper published in the *Proceedings of the 10th CAA International Computer Assisted Assessment Conference*, 4/5 July 2006, edited by Myles Danson, ISBN 0-9539572-501.

Integrated STEM in Elementary Grades Using Distributed Agent-based Computation

Pratim Sengupta, Gokul Krishnan and Mason Wright
Mind, Matter & Media Lab, Vanderbilt University, Peabody College, Nashville, TN, U.S.A.
{pratim.sengupta, gokul.krishnan, mason.wright}@vanderbilt.edu

Keywords: Computational Thinking, Tangible Interfaces, Agent-based Programming, Visual Programming, K-12 Science Education, User-centered Design, Integrated STEM, Engineering Education.

Abstract: We investigate how the integration of visual agent-based programming and computationally augmented physical structures can support curricular integration across STEM domains for elementary grade students. We introduce ViMAP-Tangible, a socio-technically distributed computational learning environment, which integrates ultrasonic sensors with the ViMAP visual programming language using a distributed computation infrastructure. In this paper, we report a study in which 3rd and 4th grade students used ViMAP-Tangible to engage in collaborative design-based activities in order to invent "drawing machines" for generating geometric shapes. The curricular activities integrate engineering practices such as user-centered design, mathematical reasoning about multiplication, rates and fractions, and physical science concepts central to learning Newtonian mechanics. We identify the key affordances of the learning environment and our pedagogical approach in terms of the relationship between the structural elements of students' physical constructions and computational models, and their learning outcomes, both in terms of computational thinking, and the domain-specific, mathematical and scientific knowledge that they began developing.

1 INTRODUCTION

Integration of the individual domains of science, technology, engineering, and mathematics (STEM) is now recognized as a central pedagogical aim of engineering and science education reform at the K-12 level (Nathan, Srisurichan, Walkington, Wolfgram, Williams, & Alibali, 2013; Berland, 2013). STEM integration is considered in the US education policy statements to be necessary for several objectives: a) supporting STEM education, including the preparation of future STEM researchers; b) for developing informed citizens; and, c) for supporting workforce development in an increasingly complex economy (Katehi et al., 2009; NRC, 2007, 2010; Nathan et al., 2013).

Integrated STEM necessitates integrating diverse domains by highlighting *big ideas* that transcend these different domains (Nathan et al., 2013; Schunn, 2009; Roehrig, Moore, Wang, & Park, 2012). However, the generation of these ideas involves material agency, conceptual agency and participation in a community of practice. For example, engineering educators have argued that some of the most essential "skills" in engineering "arise out of engagements not only with formal representations, but also with tools, materials, and other people" (Johri & Olds, 2011, p. 163). Similarly, historians, philosophers and sociologists of science have shown that the development of scientific knowledge (e.g., *big ideas* such as laws of physics) is deeply intertwined with the invention of representational systems and tools, as well as the development of communicative representational practices (e.g., modeling) (Giere, 1999; Pickering, 1993). The representational systems and tools include both semiotic systems (e.g., calculus and computational modeling languages) and mechanical devices (e.g., bubble chambers and particle accelerators). This is known as the *Science as Practice* perspective, and has been adopted as a key pedagogical framework for K-12 science education in the US (NRC, 2008).

Given this background, we believe that the focus on Integrated STEM is synergistic with the recent focus on computational thinking (Wing, 2006, 2010; NRC, 2010). Wing (2006) described computational thinking as a general, analytic approach to problem solving, designing systems, and understanding human behaviors. Sengupta, Kinnebrew, Basu,

Biswas & Clark (2013) argued that computational *thinking* is evident in the form of *epistemic and representational practices* such as problem solving, design, programming, and modeling.

Pedagogy that supports the development of representational practices associated with computational thinking can bring together different domains in science, such as biology and physics, in middle school classrooms through the use of agent-based, visual programming languages designed specifically for modeling scientific phenomena (Sengupta et al., 2013). In agent-based programming, a user creates a computer program by using simple rules to command the movement and behavior of computational agents, e.g., the Logo turtle (Papert, 1980; diSessa, Abelson & Polger, 1991; Repenning, 1993; Kelleher & Pausch, 2005; Resnick et al., 2009; Sengupta & Farris, 2012; Sengupta et al., 2013). In this paper, we extend this argument and propose that a particular form of agent-based programming and modeling, in which control of a computational agent is socio-technically distributed, can be leveraged to integrate diverse STEM domains for children in elementary grades.

The learning environment we present here can be best understood as a *socio-technically distributed* activity system. This is because ViMAP-Tangible distributes the control of a single computational agent *socially* between two collaborating users, and *technologically*, between two physical machines and a virtual algorithm. The goal of each student dyad in this study was to design, build, test and refine a hybrid (computational and physical) computational machine for generating geometric shapes. Our pedagogical approach emphasized User-Centered Design (UCD), i.e., students were asked to design their machines with *usability* (Norman, 1998) as a key focus, which has been shown to be a crucial element of product engineering but challenging to implement pedagogically.

Our paper makes three contributions. First, we present a pedagogical framework for integrating key engineering practices – UCD, collaboration and computational thinking - with math and science education. To this end, we present a theoretical framework that integrates relevant literature from multiple domains: computational thinking, Integrated STEM, Agent-based and Tangible Computation in education, User Centered Design, Design-based Learning and Collaboration in engineering practice and education. Second, we present a technological innovation in the form of ViMAP-Tangible, and a set of curricular activities, which were designed to bring about such integration. Third, while recent efforts have focused on designing and implementing Integrated STEM curricula at the college level (Sanders, 2009), middle school (Berland, 2013) and high schools (Nathan et al., 2013), we demonstrate that younger children (3rd and 4th graders) can be brought into the fold of Integrated STEM education that also includes a focus on developing computational thinking.

2 THEORETICAL FRAMEWORK

2.1 Computational Thinking in K-12

Computational thinking is an increasingly ubiquitous epistemic and representational practice in all fields of science and engineering (Wing, 2006; NRC, 2010). As Sengupta *et al.* (2013) pointed out, computational thinking draws on concepts that are fundamental to computing and computer science, but also includes practices (e.g., modeling, abstraction, reformulation, simulation, verification) that are central to a large number of scientific, engineering, and mathematical disciplines. This sentiment is also reflected in the model ACM K-12 computer science curricula for middle schools (Tucker et al., 2003), and the recently concluded National Academy of Education panel on computational thinking (NRC, 2010), which argued for integrating computational thinking with existing K-12 curricula in other domains such as mathematics and science.

Wing (2006) argued that a key characteristic of computational thinking is design-based thinking. Design is a form of problem solving in which thinking, tool manipulation, and materials are reflected in the iterative construction of an artifact (Bucciarelli, 1994; Simon, 1969; Perkins, 1986). From a pedagogical perspective, design challenges provide learners opportunities for testing and revising their developing conceptions and understanding, and interweave action and development with reflection and refinement to facilitate deep learning (Kolodner et al., 2003). Researchers have shown that students' construction failures, when scaffolded appropriately, provide additional opportunities for learning (Kolodner et al., 2003; Papert, 1980; Harel, 1990; Penner, Lehrer & Schauble, 1998).

2.2 Integrated STEM as Pedagogy

Integrated STEM has been defined as technological and engineering design-based learning approaches that intentionally integrate concepts and practices of

science and/or mathematics education with content and process of technology and/or engineering education (Sanders, 2009; McCulloch & Ernst, 2013).

As Nathan *et al.* (2013) pointed out, integration necessitates a pedagogical approach in which fields are *integrated*, rather than merely combined (cf. Dyer, Reed, & Berry, 2006; Satchwell & Loepp, 2002). The emphasis on "integration" implies that diverse fields of knowledge and practice should be merged in a manner that reveals big ideas that transcend specific disciplines (Nathan et al., 2013; Schunn, 2009; Roehrig, Moore, Wang, & Park, 2012).

Nathan *et al.* (2013) argued that big ideas that are regarded as invariants in math and science are represented using different inscriptions, both material and semiotic, as well as in different social interactions, such as in lectures or group work. As such, Nathan *et al*. (2013) argued that STEM integration in a high school level project-based engineering classroom can be viewed as the production and maintenance of *cohesion* of invariant relations across the broad range of representations that exist in the engineering classroom.

Similarly, Berland also argued that STEM integration can be brought about by a particular class of activities that she termed *STEM-design challenges* (Berland, 2013). In such activities, students are posed an engineering design challenge that can only be completed when relevant math and science concepts are applied (e.g., Coyle, Jamieson & Sommers, 1997; Fortus et al., 2004; Kanter, 2010; Kolodner et al., 2003). In such activities, pedagogically, these concepts represent domain-specific learning goals in science and math.

One of our central goals in this paper is to demonstrate that computational representational practices that are supported by agent-based programming and the design and development of usable physical control mechanisms for controlling agent-behaviors, can integrate representational practices and conceptual development across multiple STEM domains.

2.3 Agent-based Visual Programming & Tangible Computation for Children

The literature on designing agent-based programming languages and environments for novice programmers highlights the following affordances. First, agent-based programming has been shown to be intuitive for novice programmers, as it leverages learners' embodied intuitions about movement in space (Papert, 1980). Second, agent-based programming can also help children learn scientific concepts in physics and biology (diSessa, Abelson & Ploger, 1993; Repenning, 1993; Sengupta & Farris, 2012; Sengupta et al., 2013). Third, agent-based programming can help develop computational literacy through the design of self-expressive digital narratives and games using visual agent-based environments such as Scratch (Maloney et al., 2004) and Alice (Conway, 1997).

Our choice of visual programming as the *mode* of programming is grounded in the literature on computer science education. Children find it difficult to understand the syntax and semantics of programming (Spohrer & Soloway, 1986; Perkins, 1988). They also find it challenging to effectively control the flow of a program using loops and conditionals (du Boulay, 1989). Researchers have also found that alleviating syntax problems helps students focus on the semantic ones (Hohmann, 1992; Soloway, 1993; Anderson, 1989; Mannila, Peltomaki & Salakoski, 2006). Visual programming – in which students construct programs using graphical objects in a drag-and-drop interface– has been shown to be effective in alleviating these difficulties (Kelleher & Pausch, 2005). Examples of agent-based visual programming environments are AgentSheets (Repenning, 1993), StarLogo TNG (Klopfer, Yoon, & Um, 2005), Scratch (Maloney et al., 2004), ViMAP (Sengupta, Farris & Wright, 2012) and Alice (Conway, 1997).

Researchers have also started focusing in on the integration of tangible computation with agent-based modeling and programming for novice learners (Suzuki & Kato, 1995; Blackwell, 2003; Horn & Jacob, 2007; Blikstein & Wilensky, 2009). Blikstein and Wilensky (2009) showed that linking multi-agent computer models with real-world phenomena by using sensors could enable undergraduate students to learn authentic scientific and engineering concepts and practices. Tangible programming languages such as AlgoBlock (Suzuki & Kato, 1995) and Tern (Horn & Jacob, 2007; Horn et al., 2011) have also been used to teach young children programming.

In AlgoBlock, lexical elements of the Logo language are assigned to sealed metal boxes, about 20 cm on each side. These blocks can be assembled by plugging each block, via connectors on the sides, into neighboring blocks. These blocks are connected to the computer using a wired interface. A complete Logo algorithm could be constructed by assembling a sufficient number of blocks on a tabletop, and the

results of running the program are visible in the form of an animated submarine on screen (Suzuki & Kato, 1995). Horn and colleagues developed Tern, in which users construct programs by arranging and organizing wooden blocks with computer-vision fiducials (black and white symbols), which are then scanned by a program to generate a Logo algorithm. Tern has been used effectively to teach young learners programming and robotics, both in informal and formal settings (Horn, Crouser & Bers, 2011; Horn & Jacob, 2007).

While tangible and visual programming offer two different interactional modes of programming for the learner, Horn, Crouser & Bers (2011) have argued for introducing a hybrid approach in which learners (users) can elect to work either using tangible blocks, or by using a visual (graphical) programming interface, to generate the same algorithm. They found that such a hybrid approach was more advantageous than using either approach individually.

2.4 User Centered Design: Practice and Pedagogy

User Centered Design (Norman, 1998; Norman & Draper, 1986) emphasizes the importance of understanding the needs of the users in order to design usable systems. Norman (1998) argued that central principles of designing for people to support understandability and usability are: a) providing a good conceptual model to the user, which will allow them to predict the effects of their actions on the designed system; and b) making things visible, i.e., by carefully considering the relationship between the design of "controls" of a system and its "placement" (location) so that it makes the function of the control intuitively available for the user.

However, there is variability in how design researchers have defined and operationalized UCD: while there is a general agreement that UCD requires paying attention to the needs of the user, and involving the user in the system design process, there is relatively less agreement on how user involvement can be accomplished (Vredenburg, Mao, Smith, & Carey, 2002; Gulliksen, Göransson, Boivie, Blomkvist, Persson, & Cajander, 2003; Karat, 1996). Vredenburg *et al.* (2002) defined UCD as the practice of the following principles: the active involvement of users for a clear understanding of user and task requirements, iterative design and evaluation, and a multi-disciplinary approach. Gulliksen *et al*. (2003) further elaborated on the principles and argued that in addition to involving the user throughout the design process, the following principles are necessary to be enacted in practice in order to support UCD: 1) rapid prototyping during the early phases of the design process; 2) a cyclic iterative process of designing solutions interwoven with evaluation; 3) multi-disciplinary design teams that bring together distributed expertise for the various design components; and 4) an integrated design process, in which the system, the work practices, on-line help, training, organization, etc. should be developed in parallel.

In this paper, we adopt Gulliksen *et al.*'s (2003) definition of UCD as a key element of our pedagogical approach. There is some evidence that designing for instructional use can act as a productive pedagogical model for K-12 science education (Carver et al., 1992; Harel, 1990; Brown and Campione, 1993). Brown and Campione (1993) noted that fifth and sixth graders developed a deeper understanding of science concepts while they sought causal explanations to incorporate into HyperCard documents they developed to teach their classmates. A few studies have also focused on children designing agent-based instructional software for mathematics (e.g., Harel, 1990) and instructional games in science and mathematics (e.g., Kafai *et al.*, 1998). In all of these studies, students not only developed a deeper understanding of the target science or math concepts, but also developed substantial expertise in programming.

An interesting finding across these studies is that children find the consequentiality of their design projects in terms of designing for use to be quite motivating, but at the same time, they do not regard the involvement of users as a useful component of their design process (Carver et al., 1992; Kafai et al., 1998). Carver *et al*. found that "getting someone to try out the presentation" was regarded by middle school students as one of the least important tasks to accomplish during their design process; instead, they believed that the designers themselves could act as users during the design process. Similarly, in Kafai *et al*.'s work, children who designed educational software did so largely without involving real users in their design process. This in turn resulted in the design of user interfaces that were confusing for the real users (Kafai et al., 1998). This is in striking contrast to Norman's (1998) famous dictum that because the interface guides the interactions between the user and the product, it should therefore guide the design of the rest of the product.

These studies suggest that the involvement of users during the design process therefore requires explicit instructional scaffolding by the teacher. In

fact, once "real" users tested the children's designs, Carver *et al.* (1992) found that the design documents designed by the children to scaffold user interaction with their designs increased greatly in terms of making explicit the connections between the different aspects of their design, as well as explaining how to use the designed artifact. In Norman's terms (Norman, 1998), one can therefore conclude that involvement of user feedback during the design cycle enhances the usability of the product, as it makes clear to the designers the need to *make things visible* to the user.

2.5 Collaboration in Engineering Practice & STEM Education

Collaboration and teamwork hold a significant place in engineering practice (Bucciarelli, 1994). As Johri (2012) pointed out, engineers increasingly work collaboratively around the globe; technology is a primary driver of such arrangements. Anderson, Courter, McGlamery, Nathans-Kelly and Nicometo (2010) studied engineering work across six engineering firms that examined engineers' self-valuation of important work characteristics. They found remarkable similarities across settings in which most engineers saw their work as collaborative problem solving, and greatly valued communication skills and working in a team.

In the field of education research, collaboration and cooperation have been defined as conceptually distinct processes. Roschelle & Teasley (1995) defined collaboration as "a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" (Roschelle & Teasley, 1995; p. 70). Dillenbourg (1999) defined the distinction between collaboration and cooperation as follows: "In cooperation, partners split the work, solve sub-tasks individually and then assemble the partial results into the final output. In collaboration, partners do the work 'together.'" (p. 8).

Productive collaborative tasks create *positive interdependence* among learners, which can be understood as the coordinated activity of multiple people for accomplishing a specific learning objective (Kreijns, Kirschner & Jochems, 2003; Antle & Wise, 2013). White & Pea (2011) argued that collaborative tasks are likely to be most effective when they are sufficiently open-ended and complex to necessitate contributions from each member of a group (e.g., see Cohen, 1994), and when participants engage the task and one another in ways that sustain their diverse contributions (e.g., see Barron, 2003). In the literature in computer supported collaborative learning, collaboration is often fostered through variations on the "jigsaw" script in which each student only has access to part of the information (i.e. one piece of the puzzle) needed to solve a collaborative task (Miyake, Masukawa & Shirouzou, 2001).

Research suggests that collaboration can indeed enhance science and math learning through the creation of productive opportunities for shared inquiry and discourse. Pathak, Kim, Jacobson & Zhang (2010) showed that dyadic discourse in an ill-structured inquiry activity when exploring a scientific phenomenon using agent-based simulations, creates opportunities for reflective reasoning. Specifically in the context of learning kinematics using agent-based programming, researchers have showed that collaborative design of graphical, mathematical and computational representations of motion creates productive opportunities for within-group "conversations for conceptual change" (diSessa *et al*., 1991; Sherin et al., 1993). Similarly, in math education, researchers have focused on how learning in collaborative settings leads students to develop mathematical discourse by productively appropriating their group members' ways of talking or acting (Carlsen, 2010; Moschkovich, 2004; Radford, 2006; Lai & White, 2011).

Of particular relevance to our paper is the argument that meditational *tools* – that include both tangible and computational artifacts – can support collaborative learning by creating opportunities for each group member to attend to what the other is doing, by making actions visible and gaze observable in supporting collaborative meaning-making (Antle & Wise, 2013; Fernaeus & Tholander, 2006; Hornecker, 2005; Suzuki & Kato, 1995; Baker et al., 1999; Suthers et al., 2008). These studies showed that the presence of tangible artefacts in a shared *transaction space* (Hornecker, 2005) grounds the interaction between group members by providing a referential anchor for conversation, which can be referred to by using both verbal and gestural communication channels. Of direct relevance to our study, Fernaeus & Tholander (2006) found that when students worked collaboratively using a tangible programming language for learning agent-based programming, they formed subgroups dynamically, and furthermore, and these subgroups further collaborated with one another on the different activities to accomplish their goals.

3 THE VIMAP-TANGIBLE LEARNING ENVIRONMENT

3.1 The Representational Infrastructure

Three key elements of ViMAP-Tangible are:

(1) *Agent-based visual programming*: ViMAP (Sengupta, Farris & Wright, 2012) serves as the agent-based, visual programming language for our learning environment. ViMAP consists of library of graphical programming primitives designed specifically to support learning of mathematics, kinematics and biology, a construction zone where learners can generate an algorithm by using a drag-and-drop interface that is easy to understand for children, and a NetLogo environment (Wilensky, 1999) to allow learners to visualize the results of their algorithm. Similar to Alice and Scratch, ViMAP provides a drag-and-drop interface for constructing programs that is easy to use and understand for children. Commands in ViMAP include both domain-general abstractions (e.g., loops, conditionals), as well as domain-specific commands (e.g., for controlling speed, distance and acceleration of the turtles).

(2) *Tangible and gestural representation of digital information*: In ViMAP-Tangible, students can use gestures and/or mechanical devices in order to control different variables by linking them to sensors through their ViMAP programs. The ultrasonic distance sensors we used in this study measure the distance from the sensor itself to the nearest object in its field of operation. Some of the ViMAP programming primitives were designed specifically to allow users to set the values of different agent-variables (e.g., color, speed, acceleration, pen-width, horizontal and vertical displacements) based on the sensor-readings. For example, the "set <step-size> equal-to <sensor-reading>" command sets the distance the agent on the screen travels equal to the reading of the ultrasonic sensor. That is, when a learner uses this command in her/his algorithm, moving a hand towards the sensor (i.e., closer) will cause the turtle to travel a shorter distance.

(3) *Distributed control of virtual agent*: In ViMAP-Tangible, multiple students can simultaneously control the behavior of a single agent. In the version of ViMAP-Tangible we report here, control of the behavior and attributes of a single computational agent is distributed across two ultrasonic sensors, which are connected to ViMAP via an Arduino™ microcontroller (Figure 2). We believe that such a setting can foster *positive interdependence* (Antle & Wise, 2013), since it requires the coordinated action of both the group members in order to computationally implement and physically enact a successful and non-redundant control mechanism.

3.2 Curricular Activities

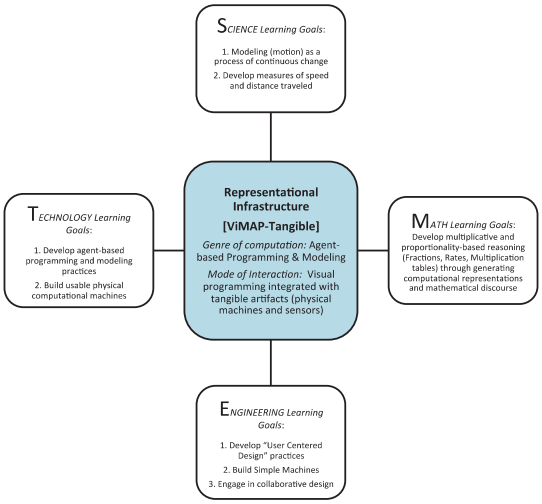


Figure 1: Domain-Specific Learning Goals for STEM Integration Using ViMAP-Tangible.

The curricular activities consisted of three phases, and were designed to integrate multiple domain-specific learning goals. These learning goals are shown in Figure 1.

During the first phase of our curriculum, students were introduced to agent-based programming using ViMAP. Students learned to generate "open" and "closed" geometric shapes (e.g. squares, circles, spirals) using ViMAP. In the second phase, students used these shapes to represent models of phenomena involving continuous change over time. We watched segments of the movie *The Lorax* in class, after which students identified various events depicted in the movie (such as Lorax running and accelerating, objects in free fall, etc.) that could be modeled using ViMAP shapes. This phase was designed to establish shape drawing as a *consequential* (Gresalfi & Ingram-Noble, 2008) activity that is not only valuable for its visual aesthetics, but that can also be used as a computational model of change over time. In both these phases, each student worked individually. Our previous work has shown that these two phases can effectively introduce students to the basics of agent-based programming, and students also learn to use their programs (i.e.,

geometric shapes) as models of motion in Phase 2 (Sengupta & Farris, 2012).

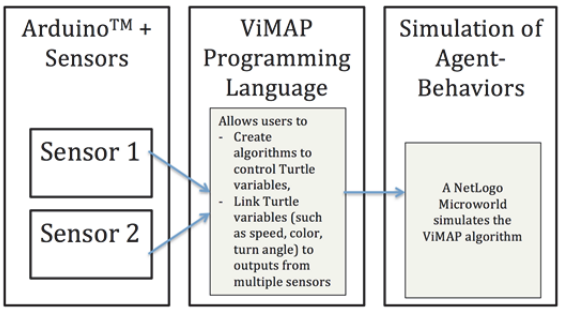


Figure 2: The ViMAP-Tangible Distributed Computation Infrastructure.

The third phase used a STEM design challenge (Berland, 2013), in which students worked in dyads to design and develop a tool for shape drawing using ViMAP. The nature of the tool that the dyads were asked to build was a hybrid computational machine consisting of a physical control structure (i.e., a simple machine), and a ViMAP program, which would together control the movement of the virtual agent (ViMAP turtle) on screen. Each student was asked to design a mechanical control for one sensor, but the dyad was responsible for jointly designing the ViMAP program. Emphasis on User Centered Design was maintained throughout this phase as follows: First, students themselves acted as the users; thereafter, students from other groups acted as users (this is the main user-testing phase for the purposes of our analysis); and during the final day of the class, students' parents were invited to visit the class to test their children's designs.

Our pedagogical approach emphasized several of Gulliksen *et al.*'s (2003) core principles of UCD. These included the following: an emphasis on rapid prototyping during the early phases of the design process; a cyclic iterative process of designing solutions interwoven with evaluation; and within group distributed expertise for different design components.

4 RESEARCH QUESTIONS

We investigate the following research questions:

1. How are students' computational and mechanical representations shaped by the following curricular foci:
   1.1. User-Centered Design?
   1.2. Collaboration between learners?
2. What is the relationship between the structural characteristics of students' physical and computational inventions and the development of their scientific and mathematical knowledge?

5 SETTING & METHOD

The study took place in a metropolitan city in the form of an enrichment program for elementary school children, conducted in a classroom on the campus of a large private university in the mid-southern USA. Classes met once a week (9:00 a.m. to 11:30 a.m.) on six consecutive Saturday mornings. Students were recruited through an online solicitation sent to the local elementary schools. None of the students in this course had any prior programming experience, and we particularly encouraged female students to apply. Students were admitted on a first-come, first-served basis. There were 16 participants, out of which eight students were in 3rd grade, and eight were in 4th grade. Five of the students were female, and three of them were in 3rd grade. The ethnic composition of the students whose work we analyzed for this paper is as follows: White (7), Asian American (4), and African American (3). Two students were absent for multiple days, and their work has not been analyzed.

The first author acted as the lead instructor for the study. Both the authors also collected data in the form of videotaped in-depth interviews with the participants, the software and hardware artifacts (i.e., ViMAP programs and physical machines) designed by the students, and field notes. The interviews were conducted while the learners were engaged in the modeling and programming activities. In some cases, the interviews ensued when the learner called upon the researcher in order to help him or her with a difficulty. In other cases, researchers conducted interviews in order to ask learners to explain their programs or models.

Similar to Carver *et al.* (1992), we believe that the role of the teacher in such a design-based classroom can be best described as using the cognitive apprenticeship framework (Collins, Brown & Newman, 1987). That is, the teacher(s) supported students' development of representational practices in some cases by explicitly *modeling* certain elements of the design practices (e.g., by acting out the ViMAP commands physically in class), and in other cases, prompting students to *reflect* on the changes they are making to their designs in terms of their affordances.

Our data consisted of student-generated artifacts, field notes, and video-recorded interviews of

students with the researchers. We analyzed this data by identifying themes and sub-themes using the double coding method (Miles & Huberman, 1994). To present our analysis, we used a case study-based approach. Our selection of cases was guided by the following criteria: *representativeness* and *typicality*. *Representativeness* implies that the selected cases should aptly represent key aspects of the instructional and learning processes. These key aspects, in turn, are defined based on the putative research question(s). *Typicality* implies that the selected case(s) should potentially represent aspects of the process of learning experienced by the majority of the student population.

We answered the research questions (RQs) as follows. In order to investigate the role of UCD and collaboration (RQ1), we compared the intermediate and final products of the students' design activities before and after user testing (during Phase 3) to check for improvements. In order to investigate the affordances of the students' designs in terms of the scientific and mathematical concepts and discourse that the students engaged in their designs (RQ2), we analyze the structure of each group's physical machines, their ViMAP program (i.e., programming commands), and their instructions for users in terms of the type(s) of mathematical measures generated by each group. The children generated these measures in order to provide explicit instructions to users for operating and understanding their machines. These measures in turn were categorized either as a) mathematical i.e., involving either multiplicative or proportionality-based reasoning, or reasoning involved in understanding geometric coordinate systems; or b) physics-based, i.e., indicative of reasoning about the measurement of speed and distance. 20% of the data reported here was blind-coded by an additional coder not involved with our study (Cohen's Kappa = 0.95).

6 FINDINGS

6.1 Pedagogical Affordances of UCD & Collaboration

Across all the groups, we found that (1) the distributed computing infrastructure resulted in positive interdependence (Kreijns et al., 2003) among students, where members of each group made coordinated, non-redundant contributions towards a common goal; and (2) a focus on User-Centered Design resulted in students refining both the software and hardware components of their programs by engaging in iterative design cycles.

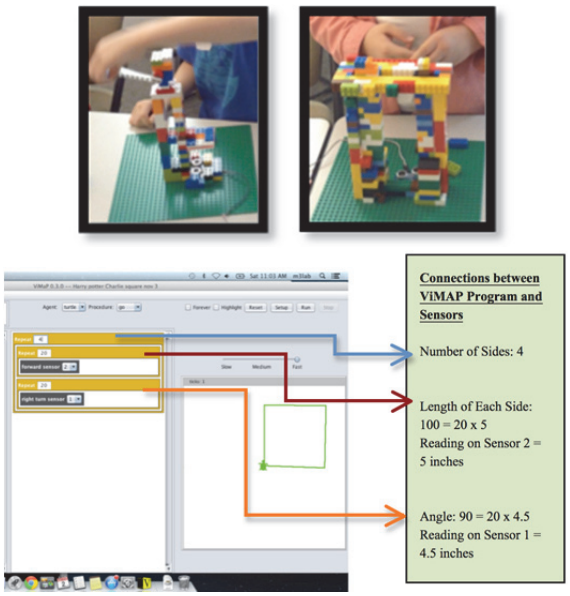


Figure 3A (Top Left): Jerry's pulley mechanism for controlling turn of the turtle via Sensor 1; Figure 3B (Top Right): Chuck's machine for controlling the speed of the turtle via Sensor 2. Figure 3C (Bottom) is a screenshot of their ViMAP program for generating a square, and shows how it responds to the sensors.

An illustrative case is the work of Chuck and Jerry (Figure 3A, 3B & 3C). During Phase 3, they built two separate machines to represent specific hand movements over the sensors. Jerry's machine consisted of a flat LEGO plate that could be lowered or raised above an ultrasonic distance sensor (Sensor 1 in Figure 3), using a pulley mechanism, to control the *turn angle* of the computational agent. Chuck's machine also comprised of a horizontal LEGO surface that could be lowered or raised above a distance sensor (Sensor 2 in Figure 3), using a manually operated crank lift, to control the *step-size* of the turtle.

In order to facilitate better collaboration between group members, during Phase 3, students' ViMAP programs became iteratively more refined with fewer bugs and redundancies. For example, in their initial versions, Chuck and Jerry had linked multiple turtle variables to each of the sensors, and after a few attempts, realized that their coordinated actions would make such a design redundant. Instead, they decided to divide their responsibilities: one person would control the turn, while the other would control the speed of the turtle, thereby, creating *positive interdependence*. Another effect of collaboration was the shared development of mathematical measures within each group, which in turn fostered

mathematical and physics-related discourse and representational practices. This is discussed in Section 6.2.

We also found that students' ViMAP programs and instructions for users, post user-testing, were more generative, and communicative. We found that the final designs of six out of seven groups allowed users to draw *multiple* shapes, whereas their initial designs during Phase 3 were more constrained and could only generate a specific shape (typically, a circle). This improvement was a direct result of user testing, as nearly every user demanded to be able to draw more than one shape using the same tool. In some cases we also noticed that the reliability of the output of their ViMAP program also improved as a result of improvement of their physical structures. For example, both Chuck and Jerry realized that they had to improve the flatness of the surfaces that were generating the sensor-readings, because their users were unable to generate reliable outputs. This resulted in Chuck introducing a flat paper strip to cover the bottom of the LEGO plates, while Jerry created a wider plate to control sensor readings more reliably.

6.2 Relationship between Children's Inventions & Learning Math and Physics

6.2.1 Designs That Primarily Supported Learning Rate and Kinematics

Figure 4 shows an example of student work in which two students, Seana and Curly, built two separate, manually operated, wheeled cars, with flat surfaces in front (stacked LEGO bricks) to represent the palm of a hand. The operating mechanism involves pushing the car towards or away from the sensor, where distance of the flat surface (representing the "palm of a hand") from the sensor generates the reading of the ultrasonic sensor. The distance of one of the cars from the sensor controls the speed of the ViMAP turtle, while distance of the other car from the second sensor controls the rotation of the turtle. A total of two groups of students developed two cars as their drawing machines.

After user-testing, Seana and Curly realized that they had to provide instructions to the user on how fast they would have to move each car in order to generate the desired shape(s). This is because, in their design, moving the two cars at different rates resulted in different shapes. Neither student had any experience in calculating rates or speed prior to this study, nor were they formally instructed during this class to calculate the rates. However, they used the computer clock and cell phones as timers, and figured out by trial and error how fast they had to move each car in terms of the time taken by their cars to travel specific distance(s) in order to generate the desired shape(s). This also resulted in making their designs more communicative by annotating the track at specific positions, along with some written and verbal instructions for users regarding how fast they needed to move the cars between these annotated positions. Although this was a rough measure of rate, we believe that this is a productive entry point into learning kinematics through the development of consequential physics talk, as well as mathematical discourse on rates – i.e., through engaging in "conversations for conceptual change" (diSessa et al., 1991).



Figure 4A: Seana and Curly's ViMAP program for shape drawing.





Figure 4B (Left) and 4C (Right): Seana (dressed in black) and Curly (dressed in blue and white) take turns in moving their "cars" towards the ultrasonic sensors. Seana controls the rotation, and Curly controls the step-size of the Turtle.

6.2.2 Designs That Primarily Supported Multiplicative and Proportionality-based Reasoning

An example of this type of design is Chuck and Jerry's work described in Section 6.1. Jerry's machine controlled the angular turn, and the user could generate readings from 1 – 10 (inches) by using the pulley mechanism. However, the user could also alter the numerical parameter of the "Repeat" command in the ViMAP program, which in turn would effectively multiply the sensor reading by that parameter. Chuck used a similar strategy to let users control the step-size of the turtle; he had created several visible marks on his towers at increments of a third of the maximum height, so that

the user can generate shapes of three levels of magnification.

Chuck and Jerry, who were both beginning to learn multiplication tables in their regular math class, thus got an opportunity to use and further develop their multiplicative and proportionality-based reasoning in order to make their designs work. We found that a total of three groups of students also invented this type of design.

6.2.3 Designs That Primarily Supported Learning the Cartesian Coordinate System

Two groups of students developed machines similar to Chuck and Jerry, but using a Cartesian coordinate system in their ViMAP program. One of their machines controlled the translational displacement of the turtle, while the other controlled the vertical displacement of the turtle. To do so, they used ViMAP commands such as "jump-X-by <Sensor-reading>" and "jump-Y-by <Sensor-reading>", respectively. An illustrative example is the case of Ken & Yang. As they iteratively refined the various combinations of translational and vertical displacements that would generate the different shapes, they became familiarized with the NetLogo XY-coordinate space. They realized during user-testing that it is much easier for users to control their machines if they are provided with coordinates for the vertices of target shapes such as equilateral triangles and rectangles. This was evident in their instructions for users on the final day.

7 CONCLUSION & DISCUSSION

In this paper, we have presented a socio-technically distributed pedagogy for Integrated STEM in elementary grades. Integrated STEM is a necessarily interdisciplinary enterprise, and given the nature of elementary classrooms (K-4), where the same teacher is responsible of teaching all disciplines, we believe that K-4 is a great setting for such curricula. The cases we presented show that 3rd and 4th grade students were indeed able to successfully participate productively in an engineering design process that highlights User Centered Design and collaboration – key elements of the practice of engineering. Furthermore, we also showed that in the process, students were able to iteratively develop expertise in agent-based programming, as well as mathematical and scientific discourse and representational practices such as measurement and modeling.

Our paper also represents the first attempt to integrate agent-based computational thinking with engineering, curricular science and math within a single curriculum in elementary grades. To this end, we presented a pedagogical innovation in the form of new socio-technically distributed computational environment - ViMAP-Tangible - and a set of curricular activities in the genre of STEM Design Challenges (Berland, 2013). The computational abstractions in the form of ViMAP programming language, as well as the associated epistemic and representational practices such as measurement and modeling - act as *invariants* that facilitate *cohesion* across the STEM domains (Nathan et al., 2013).

Our curricular design proposes a mechanism for bringing about such cohesion using a socio-technically distributed computational infrastructure. Following Berland (2013), we maintained focus on engineering practices from the beginning (Phases 1 & 2) by engaging students in design, in the form of iterative development and refinement of their computational models. Furthermore, we leveraged the synergy between tangibility and collaborative learning in a manner that introduced students to another key engineering practice (i.e., collaborative design). This in turn supported the development of scientific and mathematical knowledge by creating opportunities for reflection and positive interdependence.

ACKNOWLEDGEMENTS

Financial support was provided by the National Science Foundation (NSF Early CAREER # 1150230).

REFERENCES

Antle, A. N. and Wise, A.F. 2013. Getting down to details: Using learning theory to inform tangibles research and design for children. *Interacting with Computers*, 25, 1.

Baker, M., Hansen, T., Joiner, R., & Traum, D. 1999. The role of grounding in collaborative learning tasks. *Collaborative learning: Cognitive and Computational Approaches*, 31-63.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12(3), 307-359.

Berland, L. K. 2013. Designing for STEM Integration. *Journal of Pre-College Engineering Education Research (J-PEER)*, *3*(1), 3.

Bordogna, J., Fromm, E., Ernst, E. W. 1993. Engineering education: Innovation through integration. *Journal of Engineering Education*, *82*(1), 3-8.

Boulay, B. D., O'Shea, T., & Monk, J. 1981. The black box inside the glass box: Presenting computing concepts to novices. *International Journal of Man-Machine Studies*, *14*(3), 237-249.

Bucciarelli, L. L. 1994. *Designing engineers*. MIT.

Carlsen, M. 2010. Appropriating geometric series as a cultural tool: a study of student collaborative learning. *Educational Studies in Mathematics*, 74, 95–116.

Carver, S. M., Lehrer, R., Connell, T., & Erickson, J. (1992). Learning by hypermedia design: Issues of assessment and implementation. *Educational Psychologist*, *27*(3), 385-404.

Collins, A., Brown, J. S., & Newman, S. E. 1989. Teaching the crafts of reading, writing, and mathematics. *Cognition and instruction: Issues and agendas*, 453-494.

Conway, M. 1997. *Alice: Easy to Learn 3D Scripting for Novices*. Technical Report, School of Engineering and Applied Sciences, University of Virginia, Charlottesville, VA.

Coyle, E. J., Jamieson, L. H. & Sommers, L. S. 1997. EPICS: A model for integrating service-learning into the engineering curriculum. *Michigan Journal of Community Service Learning*, 4, 81–89.

du Boulay, B. 1989. Some difficulties of learning to program. In E. Soloway & J.C. Spohrer (Eds.), *Studying the novice programmer* (pp. 283-299).

diSessa, A. A., Abelson, H., and Ploger, D. 1991. An Overview of Boxer. *Journal of Mathematical* Behavior*, 10*(1), 3-15.

diSessa, A., Hammer, D., Sherin, B., & Kolpakowski, T. 1991. Inventing graphing: Children's meta-representational expertise. *Journal of Mathematical Behavior*, 10(2), 117-160.

Dyer, R. R., Reed, P. A., & Berry, R. 2006. Investigating the relationship between high school technology education & test scores for algebra I & geometry. *Journal of Technology Education*, 17(2), 7–17.

Fernaeus, Y., Tholander, J. 2006. Finding design qualities in a tangible programming space, in: *Proceedings of Conference on Human Factors in Computing Systems*, Montreal, Quebec, Canada. ACM Press, pp. 447-456.

Fortus, D., Dershimer, R. C., Krajcik, J., Marx, R. W. & Mamlok-Naaman, R. 2004. Design-based science and student learning. *Journal of Research in Science Teaching*, 41(10), 1081–1110.

Giere, R. N. 1999. Using models to represent reality. In *Model-based reasoning in scientific discovery* (pp. 41-57). Springer US.

Gresalfi, M., & Ingram-Goble, A. 2008. Designing for dispositions. In: *Proceedings of the 8th International conference for the learning sciences-Volume 1* (pp. 297-304). ISLS.

Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., Cajander, Å. 2003. Key principles for user-centred systems design. *Behaviour and Information Technology*, *22*(6), 397-409.

Guzdial M. 1995 Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments, 4(1). 1-44.*

Harel, I. 1990. Children as software designers: a constructionist approach for learning mathematics. *The Journal of Mathematical Behavior*, 9 (1): 4.

Hohmann, L., Guzdial, M., & Soloway, E. 1992. SODA: A computer-aided design environment for the doing and learning of software design. In *Computer Assisted Learning* (pp. 307-319). Springer Berlin Heidelberg.

Horn, M.S., Crouser, R.J., Bers, M.U. 2011. Tangible interaction and learning: The case for a hybrid approach. *Personal and Ubiquitous Computing*, 16(4), 379-389.

Horn, M.S. and Jacob, R.J.K. 2007. Designing Tangible Programming Languages for Classroom Use. In *Proc. Tangible and Embedded Interaction TEI'07*, 159-162, ACM.

Hornecker, E. 2005. A design theme for tangible interaction: Embodied facilitation, in: *Proceedings of CSCW*, Paris, France. Springer, pp. 23-44.

Johri, A., & Olds, B. M. 2011. Situated engineering learning: Bridging engineering education research and the learning sciences. *Journal of Engineering Education*, *100*(1), 151-185.

Kafai, Y. B., Franke, M., Ching, C., & Shih, J. (1998). Game design as an interactive learning environment fostering students' and teachers' mathematical inquiry. *International Journal of Computers for Mathematical Learning*, 3(2), 149–184.

Kreijns, K., Kirschner, P.A., Jochems, W. 2003. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior*, 19, pp. 335-353.

Katehi, L., Pearson, G., Feder, M. A. 2009. *Engineering in K-12 education: Understanding the status and improving the prospects*. Washington, DC: National Academies Press, 2009.

Kelleher, C. & Pausch, R. 2005. Lowering the barriers to programming: a taxonomy of programming environments and languages for novice programmers, ACM Computing Surveys, Vol. (37) 83–137.

Klopfer, E., Yoon, S. and Um, T. 2005. Teaching Complex Dynamic Systems to Young Students with StarLogo. *The Journal of Computers in Mathematics and Science Teaching; 24(2): 157-178.*

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning By Design™ into practice. *The Journal of the Learning Sciences*, *12*(4), 495-547.

Lai, K., & White, T. 2012. Exploring quadrilaterals in a small group computing environment. *Computers & Education*, *59*(3), 963-973.

Maloney, J., Burd, L., Kafai, Y., Rusk, N., Silverman, B., and Resnick, M. (2004) Scratch: A Sneak Preview. In *Proc. of Creating, Connecting, and Collaborating through Computing, 104-109.*

McCulloch, A. W., Ernst, J. V. 2012. Estuarine Ecosystems: Using T & E Signature Approaches to Support STEM Integration. *Technology and*

Engineering Teacher, *72*(3), 13-17.

Moschkovich, J. N. 2004. Appropriating mathematical practices: a case study of learning to use and explore functions through interaction with a tutor. *Educational Studies in Mathematics*, 55, 49–80.

Miles, M. B. & Huberman, A. M. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*.

Miyake, N., Masukawa, H., Shirouzou, H. 2001. The complex jigsaw as an enhancer of collaborative knowledge building in undergraduate introductory science courses. In *Proc. of European CSCL*, 454-461.

NAE and NRC. 2009. *Engineering in K–12 Education: Understanding the Status and Improving the Prospects.* The National Academies Press.

Nathan, M. J., Srisurichan, R., Walkington, C., Wolfgram, M., Williams, C., & Alibali, M. W. 2013. Building Cohesion Across Representations: A Mechanism for STEM Integration. *Journal of Engineering Education*, *102*(1), 77-116.

National Research Council (NRC). 2008. *Taking science to school: Learning and teaching science in grades K–8.* National Academy Press.

National Research Council (NRC). 2010. *Report of a Workshop on The Scope and Nature of Computational Thinking.* The National Academies Press.

Norman, D. A. 1998. The design of everyday things. New York.

Norman, D. A., & Draper, S. W. 1986. *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc.

Papert, S. 1980. Mindstorms: children, computers, and powerful ideas. Basic Books, Inc. New York, NY.

Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1997). Building functional models: Designing an elbow. *Journal of Research in Science Teaching*, *34*(2), 125-143.

Perkins, D. N. 1986. *Knowledge as design*. Hillsdale, NJ: L. Erlbaum Associates.

Perkins, D. N., & Simmons, R. 1988. Patterns of misunderstanding: An integrative model for science, math, and programming. *Review of Educational Research*, *58*(3), 303-326.

Pickering, A. 1993. The mangle of practice: Agency and emergence in the sociology of science. *American Journal of Sociology*, 559-589.

Redish, E. F. and Wilson, J. M. 1993. Student programming in the introductory physics course: M.U.P.P.E.T. *Am. J. Phys.* 61: 222–232.

Repenning, A. 1993. Agentsheets: A tool for building domain-oriented visual programming, *Conference on Human Factors in Computing Systems, 142-143.*

Roschelle, J., & Teasley, S. D. 1995. The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Springer.

Roehrig, G. H., Moore, T. J., Wang, H. H., & Park, M. S. 2012. Is Adding the E Enough? Investigating the Impact of K-12 Engineering Standards on the Implementation of STEM Integration. *School Science and Mathematics*, *112*(1), 31-44.

Sanders, M. 2009. Integrative STEM education: Primer. *The Technology Teacher*, 68(4), 20–27.

Satchwell, R. E., & Loepp, F. L. 2002. Designing and implementing an integrated mathematics, science, and technology curriculum for the middle school. *Journal of Industrial Teacher Education*, 39(3), 41–66.

Sengupta, P., & Farris, A. V. 2012. Learning kinematics in elementary grades using agent-based computational modeling: a visual programming-based approach. In *Proceedings of 11th International Conference of Interaction Design & Children* (pp. 78-87). ACM.

Sengupta, P., Farris, A.V, & Wright, M. 2012. From Agents to Aggregation via Aesthetics: Learning Mechanics with Visual Agent-based Computational Modelling. *Technology, Knowledge & Learning*. 17 (1-2), pp 23 - 42.

Sengupta, P., Kinnebrew, J.S., Basu, S., Biswas, G., & Clark, D. 2013. Integrating Computational Thinking with K-12 Science Education Using Agent-based Computation: A Theoretical Framework. *Education and Information Technologies,* 18 (2), 351-380.

Sherin, B., diSessa, A. A., & Hammer, D. M. 1993. Dynaturtle revisited: Learning physics through collaborative design of a computer model. *Interactive Learning Environments*, 3(2), 91-118.

Simon, H. A. (1969). *The sciences of the artificial.* Cambridge, MA.

Spohrer, J. C., & Soloway, E. 1986. Novice mistakes: Are the folk wisdoms correct? *Communications of the ACM*, *29*(7), 624-632.

Suzuki, H., Kato, H. 1995. Interaction-level support for collaborative learning: AlgoBlock - an open programming language, in: *Proceedings of CSCL*, Erlbaum Associates, Hillsdale, NJ, pp. 349- 355.

Tucker, A., Deek, F., Jones, J., McCowan, D., Stephenson, C., & Verno, A. 2003. A model curriculum for K-12 computer science: Final report of the ACM K-12 task force curriculum committee. *New York, NY: ACM*.

Wilensky, U. 1999. *NetLogo* [Computer Software]. Northwestern University, Evanston, IL.

Wing, J. M. 2006. Computational Thinking. *Communications of the ACM,* vol. 49(3), pp. 33-35.

SHORT PAPERS

Layered Evaluation Based on Timed Automata for Educational Scenarios
Case Study on Blended Learning

Manel Bensassi and Mona Laroussi
RIADI Laboratory, School of Engineering in Computing Science (ENSI), University of Manouba, Manouba, Tunisia
bensassi.manel@gnet.tn, mona.laroussi@planet.tn

Keywords: Educational Scenario, Blended Learning, Evaluation, Formal Verification, Learning Environment, Timed Automata, Fractal.

Abstract: The evaluation of educational scenario is important to educators. It is one of the most challenging tasks that the pedagogical designer is faced with. In fact, the high cost of the implementation of learning scenarios in distance learning and the complexity of their context and their functionalities emphasize the need to evaluate and to simulate them in order to detect inconsistencies and incompleteness and to later provide valuable feedback and recommendations to designers. In this paper, we propose a formal and layered evaluation based on Timed Automata theory. We believe that this approach allows designers to evaluate scenarios at an early stage to avoid costly and time-consuming scenarios implementation or deployment. The proposed approach aims to simulate scenario's execution on real time in order to detect deficiencies (such as deadlocks and livelocks) and then to regulate scenario with contextual constraints. We outline how this evaluation could be applied at different levels of learning scenarios' granularity.

1 INTRODUCTION

Educational scenarios and learning objects are growing in scale and functionality. Giving this large set of existing techniques and a practical reuse and share of this learning database, evaluation of learning scenarios and systems is becoming more important than inventing new techniques (Brusilovsky et al., 2004). In fact, the high cost of the implementation of educational scenarios in distance learning and the complexity of their context make the evaluation step essential to understand whether a particular technique works as expected and in which context or application areas it works the best (Brusilovsky et al., 2004).

Much research on this area has focused on exploring the dimensions of evaluation and applies it to evaluate the quality of the educational scenario (Bubaš et al., 2007). Other works focus on evaluating the workflow in the courseware (de Moor, 2007). In these methods, functionalities of the online course are scored on both their effectiveness and importance for the activities they support. The activities themselves are also scored on their importance to the organization. The important scores act as weight for the effectiveness scores. These methods evaluate the learning scenario "*as a whole*".

In order to remedy to this lack, a number of researchers recommend evaluating adaptive systems on a layer-by-layer basis as a more comprehensive approach (Brusilovsky et al., 2004)(Paramythis et al., 2010) based on user modelling analysis.

The review of learning scenario evaluation methods shows that the lack of collected data, as well as the difficulty in their generalization constitute a major barrier to evaluating process (Brusilovsky et al., 2004) In fact, designers find difficulties in getting information about student and their learning behavior Besides, these traditional approaches of evaluation, although useful in theory and in practice, could be applied *only after the execution* of the learning scenario. These approaches require that the whole system had to be developed and deployed. Thus, designers couldn't anticipate and predict learners' difficulties. We argue that it is essentially important for each new or existing learning scenarios and systems to be properly evaluated (Brusilovsky et al., 2004).

At the same time all of these approaches do not take into account the different phases, processes and components of the learning system (Brusilovsky et al., 2004). Finally, traditional evaluation provides no feedback about performance of different system components. Thus, successful design practices cannot be easily re-used across different applications and services (Brusilovsky et al., 2004).

As a solution to the listed problems, we propose, in this paper, an early formal evaluation method able to model educational scenarios at different levels of granularity: activity, scenario and system. Our approach relies on timed automata and tracks simulation. It takes into account pedagogical and contextual requirements.

This paper is organized as follows. In section two, we discuss the benefit of fractal educational scenario. We outline how this abstract concept can improve the current evaluation practice and our understanding of learning scenario. Section three focuses on our evaluation methodology definition. In order to demonstrate the benefits of the proposed approach, we present, in section four, a realistic example. In section five, we discuss the evaluation of this approach. We conclude this article with perspective and future works.

2 FRACTALITY IN LEARNING SCENARIO

Our purpose in this paper is to present a generic and a formal model of evaluation of learning scenario. This model helps designers to understand and analyze the fine and coarse grained educational scenario regardless of the learning approaches (distance, pervasive, mobile, face to face, blended ...). Our methodology consists in decomposing the course ware and to define it as fractal-activities.

2.1 Fractal Educational Scenario Definition

The expected path through the online course is given by planned activities. This allows data to be aggregated at different levels of granularity to stimulate collaboration, problem-solving and learning as well as access to learning resources. The hierarchical activities' correlation leads to a coherent and collective scenario, that we call here "*Fractal Educational Scenario*". It is an abstract model that has got a major dimension: *"the activity"*/

The fractal learning scenario model may be used and reused by other educational scenarios and the composition of them may be exposed as a new educational scenario.

The fundamental feature of fractal educational scenario is the separation between the description (which represents the scenario's interface) and his implementation and deployment. This implementation of learning environment and services can evolve overtime without disturbing the learner or the tutors of the educational scenario. The same interface can be offered by many implementations, or conversely, these implementations could change without affecting the aggregate scenario (synchronous activities like chat room keep their invariant description). Then, this educational scenario model is fractal and the newly-formed scenario is an educational scenario itself.

The learning activity is the major component in educational scenarios regardless of the learning approach. Consequently, it represents the invariant scale that does not change with the change of the scale (granularity) of the fractal educational scenario. Then, we consider these educational scenarios fractal-activities. A learning scenario is composed of one or more activities that are bounded through semantic links (requirement, objectives, prerequisite) and temporal links (before, after, overalaps,...).

2.2 Challenges of Fractal Educational Scenario

The use of fractal approach in learning scenario's definition offers some specific benefits:

- *It is an abstract concept* of learning scenario expressed independent of any design model (UML, IMS LD...) and learning approaches (mobile, blended, pervasive,..).
- *It's an incremental scenario*. Designers shoudn't be required to dramatically reengineer their educational scenario to obtain conceptual and learning satisfaction. Rather, they should be able to leverage existing educational scenarios in doing so.
- *Agility* in sharing, reusing and evaluating is often promoted as an important benefit of fractal educational scenario which is an organization with a workflow implemented on a loosely coupled infrastructure.

3 LAYERED EVALUATION BASED ON TIMED AUTOMATA MODELING AND SIMULATION

3.1 Layers of Evaluation

Based on the concept of fractality, we decompose the learning scenario into different and dependent layers. In order to facilitate the evaluation practice, we propose to evaluate the learning scenario layer-by-layer (shown in figure 1):

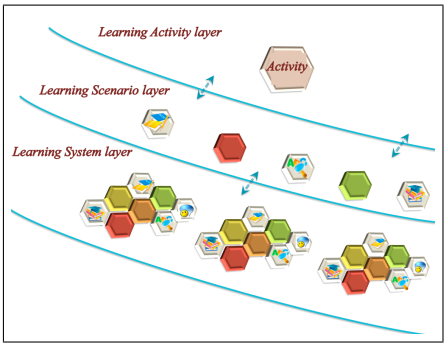


Figure 1: Evaluation of different Layers of fractal educational scenario.

- *The lower layer: Evaluating the activity* that represents the basic and the finest granular of the learning scenario. At this layer, each activity is being evaluated separately.
- *The medium layer: Evaluating the workflow between activities* that compose the scenario.
- *The higher layer: Evaluating the learning system* At this layer, we evaluate the correlation between different scenarios that compose the learning system.

We believe that the proposed framework provides an insight into the design of each activity and phase of the learning process and thus facilitating the improvement of conceived learning scenarios. The decomposition of the learning scenario into layers contributes to the reuse of each learning components.

3.2 Evaluation Stage

Evaluation process is of a great importance in all aspects and stages of teaching and learning. Indeed, it should supervise the learning experience to be both more efficient and effective as the feedback is used to continuously improve matters. There are two possible stages for educational scenario evaluation:

1. The evaluation of the conceived educational scenario before implementation and deployment
2. The evaluation of the executed educational scenario.

We suggest that the evaluation of the educational scenario takes place at an early stage in the life cycle. This will tend to detect learner's difficulties. It will allow us to focus inwardly on various constraints and factors (such as activities' timing, learners' interaction, educational scenario setting, implementation strategies..).

3.3 Evaluation Process

In this section we present our proposed evaluation process that is based on timed automata modeling and formal verification. It contains essentially these steps:

3.3.1 Formal Modeling for Fractal Educational Scenario

The automata modeling consists of drawing states (represented by circles), and transitions (represented by arrows). An automaton contains a finite set of states. At each instant in time of some run, the automaton is in one of its states (Alur and Dill, 1994).

Here, we consider educational scenario as a transition system built on a set of states containing useful information like the activities' status. Thus, an automaton is composed of two parts: the states part for describing learning activities and the transition function part for representing the advancing over activities. An automaton contains a finite set of states. At each instant in time of some run, the automaton is in one of its states. At each time step, when the automaton reads a symbol, it jumps or transits to a next state that is decided by a function that takes current state and the symbol currently read as parameters. This function is called transition function.
Each activity will be modeled as an automaton. The whole scenario will be the parallel composition of different activities that will be a synchronized switch to the following temporal domain.

Temporal Domain Definition. We define activity 's extensible phase, noted P_A , as being a sequence of ordered moments $\{t\}$ or a time intervals that are marked by a beginning and an ending moment $\lfloor t^-, t^+ \rfloor$.

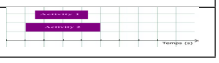
$$(\exists t^-)(\forall t \in P_A)(t^- \leq t) and (\exists t^+)(\forall t \in P_A)(t \leq t^+)$$

We note that t^- and t^+ are the limit points for the activity's achievement. Given the activity's phase specification, we define the temporal domain of educational scenario as the union of the different activities' phases forming the online course. We refer to Allen's relation so as to specify Activities' temporal connection (table 1).

3.3.2 Simulation and Verification

Using automata theory as a method of specification, it is possible to use the formal verification techniques to demonstrate that a system design is correct with respect to its specification. The main objective of formal specification and verification is to minimize the creation of functional errors and limits. This analysis will usually reveal errors or consistencies through

Table 1: Temporal Domain Definition.

Description	Notation	Formal definition	Figure
Simultaneous' activities	$P_{a1} = P_{a2}$	$(P^-_{a1} = P^-_{a2}) \wedge (P^+_{a1} = P^+_{a2})$	
Activity 1 before Activity 2	$P_{a1} < P_{a2}$	$(P^+_{a1} < P^-_{a2})$	
Activity 1 overlaps Activity 2	$P_{a1} \underline{o} P_{a2}$	$(P^-_{a1} < P^-_{a2}) \wedge (P^+_{a1} > P^-_{a2}) \wedge (P^+_{a1} < P^+_{a2})$	
Activity 1 during Activity 2	$P_{a1} \underline{d} P_{a2}$	$(P^+_{a1} > P^-_{a2}) \wedge (P^+_{a1} < P^+_{a2})$	

simulations.

In our context, formal specification methods can solve many problems (track's collecting, monitoring problems, behavioral study of learners..), guaranteeing non-ambiguity and supporting powerful analysis capabilities. This process aims to build a remedial scenario and to help designer in the reengineering process by providing errors and warning reports. In the scope of our study, we tried to check the following properties:

1. **Deadlock** is a situation in which two or more active entities (for example students or groups) are each waiting for the other to finish, and thus neither ever does.
2. **Liveness** is similar to a deadlock, except that the states of the processes involved in the livelock constantly change with regard to one another, none progressing.
3. **Starvation** describes a situation where a subgroup is unable to gain regular access to shared resources (for example help of the coach) and is unable to make progress. This happens when shared resources are made unavailable for long periods by "greedy" actors. We want to make sure that each Subgroup will do his activity in time.

In the following section, we are going to show how we can apply this formalism at the educational scenario and use it to predict student's challenges.

4 CASE STUDY OF BLENDED LEARNING SCENARIO

We want to show that automata formalism can describe the different learning activities and all the paths to follow in order to evaluate the educational scenario through simulation and tests generation. To illustrate feasibility, we select the blended educational scenario which combines classroom and online education. This scenario, described by the following activity diagram (see figure 2), is opted by a high school that decides to implement distance monitoring of students in chemical master degree. The physical setting of this trial, where activities take place, is the school's laboratory that is open for this class from 8:00 am to 12:00 am.

In order to boost intra group competition, students were divided in two groups under the supervision of their coach. Each group consisted of four students. Additionally, each group was divided in two subgroups of two students, where one subgroup was working on the first Experiment, while the other subgroup was working on the second. The ultimate goal behind this clustering is to reinforce teamwork and collaboration within the individual subgroup. Each group was required to create a learning experience utilizing various technologies.

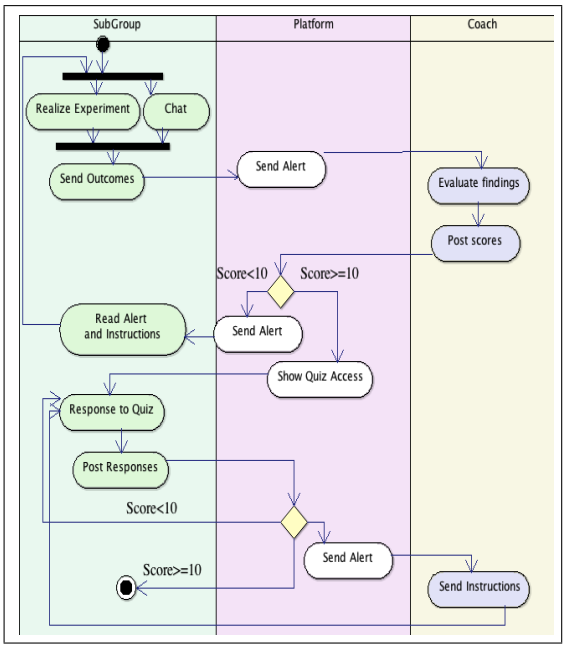


Figure 2: Initial Educational Scenario.

The groups are conducting this activity together which makes the process quite collaborative and challenging consisting of two majors stages as follows:

- Each experiment is divided into two-sub experi-

ments (see the table 2). In the first stage, students realized the first experiment.

- The second stage consists in answering to the quiz questions.

Each group is equipped with a computer and an Internet connection. At the beginning of each stage, the subgroup has got an access to the manual of experiment in the distance platform Moodle. This manual describes the activities to be accomplished. To achieve those activities, the two subgroups should collaborate together to write a common report describing their works and outcomes. They can discuss their findings through a chat room.
After a predefined time of each stage, each group must send their result to the coach to be validated. If their findings were not suitable, they would have to retry again their experience. Otherwise, an adapted quiz would be visible in the platform. Students answer the different questions and submit their responses. If their score is above 10, the description of the next experiment becomes visible. Otherwise, if the answer submitted by the group is not correct, the system sends an alert to the coach informing him/her that this group requires some support. Then, the coach should send out some instructions. Students should try to answer the questions one more time. Having completed the aforementioned stages, each group should draw their conclusions and present the outcomes of their study and experiment.

For the realization of this practical scenario, the laboratory's technician has to give the coach this data sheet of available quantity of required products.

4.1 Modelisation

To model automata and to generate simulations, we use the free software UPPAAL which is a graphical tool suite for modeling network of Timed automata (Bengtsson et al., 1996).
Figures 3, 5, 4 and 6 provide an overview of different automata modeled for the planned learning activities (Chat, Quiz and Experiment) and corresponding to different actors (the subgroup who plays the different activities, the e-learning platform and the coach).We define a local variable *"clock"* (named "D") that gives an idea about the duration of each activity. We also define an other global variable *"Timing"* to calculate the duration of the whole scenario. In order to implement our approach (fractal scenario), we model each activity separately. The idea is to define templates for activities that are instantiated to have a simulation of the whole scenario. The motivation for the use of templates is that the understanding, the share and the

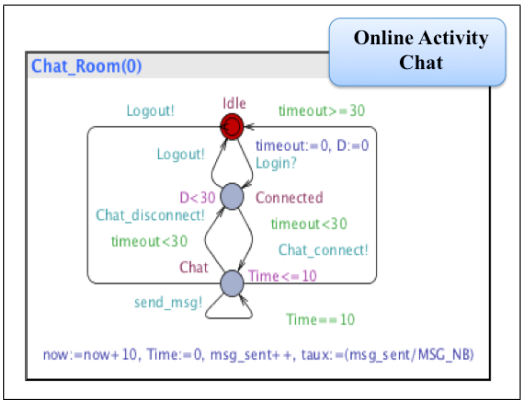


Figure 3: Overview of the automata of Chat Room activity.

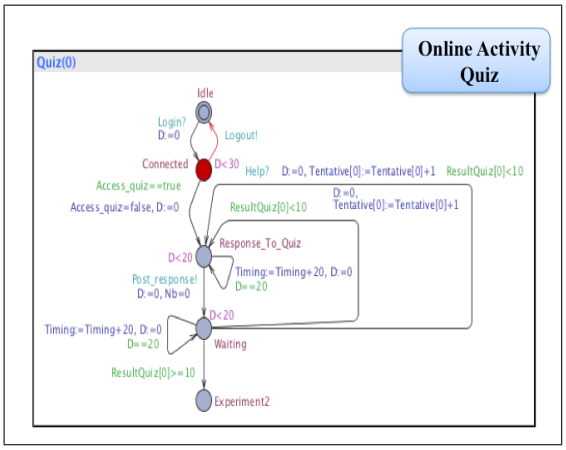


Figure 4: Overview of the automata of Quiz activity.

reuse of different components of the learning scenario become easy. We define:

1. *Chat Template* which is a finite-state machine extended with the clock variables and pedagogical indicators (collaboration rate) (see figure 3). The collaboration rate is a variable describing the number of messages written per student divided by the total number of message during the session of chat. This template serves as a model for the synchronous activity.
2. *Quiz Template* that is a finite-state machine. The edges are labelled by scoring values, and clocks constraints (the duration of the activity and the whole scenario) (see figure 4).
3. *Experiment Template*: that is a finite-state machine. The edges are labelled by a variable that monitors the quantity of product. The automata is also enriched with clocks constraints (the duration of the experiment "DExp" and the waiting "Timing") (see figure 5).

The whole scenario is modelled as a parallel composition of timed automata.

Table 2: The Datasheet of Experiments.

Group N.	Experiment N.	Duration	Resources	Required quantity (ml)	Available quantity (ml)
A1 and B1	Exp.1.1	1 H	A ⇒ A'	150	160
A1 and B1	Exp.1.2	1 H	A'+ B ⇒ A'B	200	300
A2 and B2	Exp.2.1	1 H	C ⇒ C'	100	400
A2 and B2	Exp.2.2	1 H	C' + D ⇒ C'D	50	150

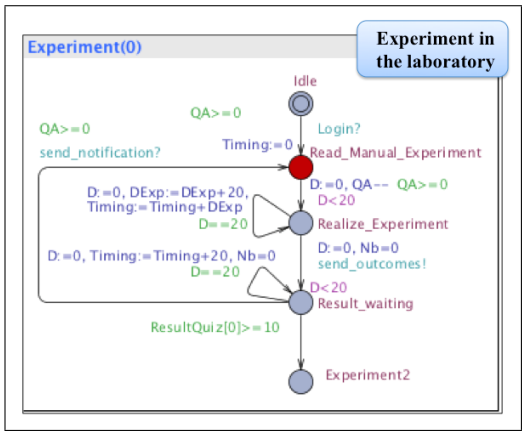


Figure 5: Overview of the automata of the face to face activity.

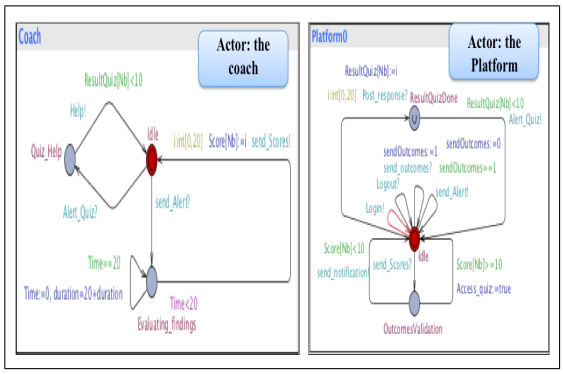


Figure 6: The automata Model of actors involved in the planned learning scenario.

An automaton may perform a transition separately or synchronize with another automaton (channel synchronization) or it can be activated after a period of time through flags. The time correlation is synchronized to switch the temporal domain previously defined. Consequently, we represent the whole scenario by automata network to obtain fractal educational scenario model.

4.2 Simulation

Based on automata presented in the previous section, tracks simulations are generated visualizing all possible interactions between different actors. A screen dump of the simulation of the designed educational scenario is below (see figure 7): Figure shows the simulator with different automata of learning activities and actors on the message sequence chart as well as the different synchronizations between the automata are shown.

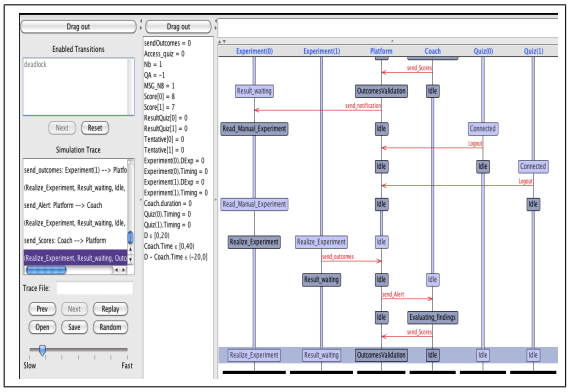


Figure 7: Simuation of the initial Educational Scenario.

In order to help designers to improve their educational scenario and to obtain better outcomes, through the generated simulations, we try to localize design errors, to answer and to verify the following questions: **(i) Does the description of the learning scenario clearly define time constraints for each activity and the whole scenario?**, **(ii) Is there any situation of deadlocks, livelocks or starvation?**

4.2.1 Layer 1: Evaluation of the Activity

1. **Experiment:** A possible situation of deadlock is detected within the tracks simulation of the activity *"Experiement"* (see figure 8). In fact, we fixe the simulation parameter and we simulate the interactoin between two subgroups automata for this activity. Different products in the laboratory are considered as shared resources. We point out that the *resource A* is available only for one iteration.
2. **Quiz:** The lack of timing constraints and the number of attempts for the quiz generate an infinite duration for this activity. Figure 9 shows that the number of attempts to response to the quiz wasn't specified so it tends to infinity.

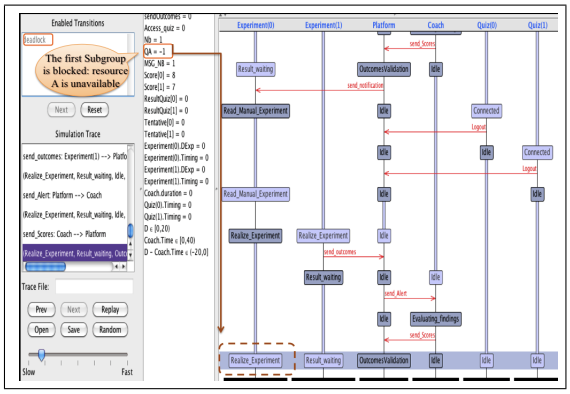


Figure 8: Resource unavailable: Student couldn't restart the experiment.

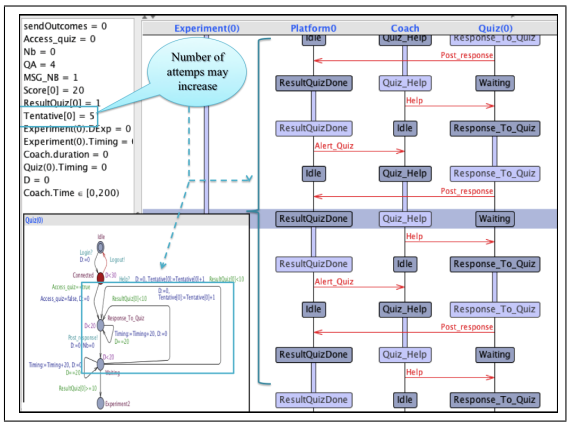


Figure 9: Student could be blocked in the quiz.

4.2.2 Layer 2: Evaluation of the Whole Scenario

Starvation: We simulate the interaction between the coach and two subgroups (see figure 10). If we consider the coach's hints as a shared resource and students ask for help at the same time. Then, we deem that one group is still inactive for a period of time, because the coach couldn't respond to all students at the same time. A situation of starvation could appear.

Following this experiment, the designer makes his decision to remedy to the proposed educational scenario and to add some changes to the initial one: he specifies the maximal number to access the quiz activity and he sets the number of hints to two.

5 RESULTS AND DISCUSSION

Several case studies have been carried out with 44 designers in the context of the international online master's program. The case studies offered feedback on the usefulness of the proposed methodology to evaluate the conceived educational scenario, the usability of the proposed approach, the opinions of designers concerning these ways of learning analysis and whether or not they would use it again.

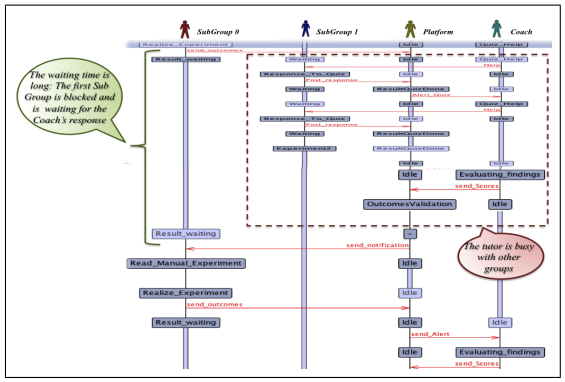


Figure 10: The Coach is strongly requested by students. He couldn't monitor a higher number of group.

In each case study, we started by asking participants questions about their level of expertise in formal verification and automata theory. This information was necessary in order to evaluate the usefulness and relevance of the transformation assistance. At the end of each experiment, the participants were called to answer a questionnaire containing 3 questions and were given the opportunity to make comments and suggestions and mention difficulties encountered.

Results and Discussion. Designers gave us their feedback and opinions on using this formal approach by answering the questionnaire. The questions and the pedagogical designers' answers are summarized below.

The first question was "*Has this approach helped you to evaluate the conceived learning scenario?*" Designers were able to choose between three options (Yes - No - Maybe). In general, designers consider that this approach helps them to evaluate the "*functional dependability*" of the learning scenario (72 percent). They explain this choice by saying that they usually have difficulties in getting proper information about learners and their learning behavior. This difficulty stems from the fact they need to collect learners' tracks and to treat them in order to get valuable feedback. With this formal approach, there is easy generation and analysis of learning tracks.

Aswering the question "*What do you think about this way of evaluation and learning analytics based on Timed Automata?*" allowed designers to choose among five options:"very interesting, interesting, makes no difference, not very interesting and uninsteresting".

Sixty-seven percent of the designers selected "very in-

teresting" and 14% selected interesting They felt that this approach allows them to evaluate the conceived learning scenario in a new and a formal way. Some designers said that the strength of this way of evaluation is that (i) it is based on a rigourous formalism to model and check the conception of the learning scenario (ii) it allows to extrapolate tracks (iii) it allows us to have a formal model representing learning activities that can be reused in different contexts. A few designers (14%) found this way of evaluation not very interesting and 5% thought it made no difference. They argued that "with this formal evaluation we couldn't evaluate the pedagogy applied and we couldn't anticipate learners' knowldge difficulties". They prefer traditional methods to evaluate the learning scenarios such as questionnaires and interviews.

A question was asked concerning the appropriateness of assistance recommendation: "*Are the assistance recommendations are appropriate?"*. Designers were able to choose between three options: yes, no and maybe. They feel that recommendations are clearly presented. In this case, the percentage of "yes" answers is 73%, while the "no" and "maybe" options were selected by 27%. Some designers say that these recommendations can be hardly understood and analyzed. Other designers commented that in some cases, recommendations couldn't be applied (financial limits).

6 CONCLUSION

This paper suggests using Timed Automata formalism in order to model and to evaluate fractal educational scenarios. This work also outlines that the evaluation of learning materials (activity, scenario and system) needs to be formal and conducted at an early stage. Then we suggest a layered evaluation framework where the success of the evaluation is addressed at different layers. This paper has attempted to demonstrate the benefits of the proposed approach by presenting a realistic case study.We think that this formal methodology of evaluation provides useful information for the re-use and the reengineering of learning materials. On the one hand, it can solve many difficulties in getting proper information about the students and their behavior (ethical problems, track s' collect,...). On the other hand, the use of a rigorous formalism to describe a model of deployment allows a better and more precise understanding of the planned scenario. It provides the designer with an analytical model that helps him detect inconsistencies and incompleteness of the designed scenario through testing case generation and deductive verification. Finally, it creates an environment that simulates an external reality and contextual constraints (time, resource, user,..).

To extend this first result further, we are currently working on two directions: Firstly, we will attempt to deepen our proposal and to apply formal verification relying on temporal logical formulas and model checking engine. Secondly, in order to assist the designer in expressing these properties, we are developping a tool that will help him/her to draw this business requirement and to generate error report automatically.

REFERENCES

Alur, R. and Dill, D. L. (1994). A theory of timed automata. *Theoretical computer science*, 126(2):183–235.

Bedell, E. F. (1984). *Computer Solution: Strategies for Success in the Information Age*. McGraw-Hill, Inc.

Bengtsson, J., Larsen, K., Larsson, F., Pettersson, P., and Yi, W. (1996). *UPPAALa tool suite for automatic verification of real-time systems*. Springer.

Brusilovsky, P., Karagiannidis, C., and Sampson, D. (2004). Layered evaluation of adaptive learning systems. *International Journal of Continuing Engineering Education and Life Long Learning*, 14(4):402–421.

Bubaš, G., Balaban, I., and Begičević, N. (2007). Evaluation of online courses as an element of instructional design: The case of two hybrid university courses. In *Conference Computers in Education, 30th Jubilee International Convention MIPRO 2007*.

de Moor, A. (2007). A practical method for courseware evaluation. In *Proceedings of the 2nd international conference on Pragmatic web*, pages 57–63. ACM.

Georgiakakis, P., Papasalouros, A., Retalis, S., Siassiakos, K., and Papaspyrou, N. (2005). Evaluating the usability of web-based learning management systems. *THEMES in Education*, 6(1):45–59.

Höök, K. (1998). Evaluating the utility and usability of an adaptive hypermedia system. *Knowledge-Based Systems*, 10(5):311–319.

Jefferies, P. and Stahl, B. C. (2005). Some ethical considerations regarding the relationship of e-learning and pedagogy. In *ETHICOMP*.

Paramythis, A., Weibelzahl, S., and Masthoff, J. (2010). Layered evaluation of interactive adaptive systems: framework and formative methods. *User Modeling and User-Adapted Interaction*, 20(5):383–453.

Weibelzahl, S. (2001). *Evaluation of adaptive systems*. Springer.

Hints for Organizing a Successful Doctoral Symposium
An Experience Report

Claudia Raibulet[1] and Eila Ovaska[2]
[1]*Dipartimento di Informatica, Sistemistica e Comunicazione, University of Milano-Bicocca, Viale Sarca, 336, Milan, Italy*
[2]*VTT Technical Research Centre, Kaitovayla, 1, Oulu, Finland*
claudia.raibulet@unimib.it, eila.ovaska@vtt.fi

Keywords: Doctoral Symposium, Phd Students, Software Engineering.

Abstract: This paper describes our joint experience in organizing a Doctoral Symposium co-located with one of the main software engineering conferences. It presents the issues we addressed during the organization of the symposium, as well as during the symposium. This paper is addressed (1) to the organizers of Doctoral Symposiums aiming to provide a recipe with the main ingredients and preparation steps and their related significance in the entire organization process, as well as (2) to the PhD students providing our feedback on the expectations, evaluation, and presentation of their contributions.

1 INTRODUCTION

PhD students in software engineering represent the pool of researchers who aim to follow an academic carrier (and, thus, teach students how to do software engineering) or to join an industrial reality (and, thus, follow the engineering of large and complex software) (Almi, 2011); (Subrahmanyam, 2009); (Villavicencio, 2012); (XueYun, 2010). In this perspective, PhD students should be able to make a research plan of high quality and to follow its implementation from all the points of view (technical and organizational/managerial). Moreover, they should be able to properly interact with all the stakeholders of their research work (e.g., in this particular context, experts in the research domain who are not the requesters/contractors of the project and do not know any detail or history of the project, but who may "buy" the idea and/or the solution).

The Doctoral Symposiums organized in the context of the main conferences give the PhD students the possibility to present their research projects (and research results) and to have a valuable feedback from experts in the software engineering domain. A presentation in the context of a Doctoral Symposium should be seen as a preview of a PhD dissertation.

This paper presents the experience of the co-chairs of a Doctoral Symposium co-located in the context of one of the most important software engineering conferences: European Conference on Software Architectures (ECSA, 2013).

1.1 Why This Paper?

The motivation beyond this paper is twofold.

The first concerns the organization of the Doctoral Symposium. The co-chairs have experience in organizing such types of events (e.g., conferences, workshops, doctoral symposiums). Concerning the Doctoral Symposiums, we have gathered quite different experiences in organizing different symposiums, differences coming from the research area and from the co-located conferences. Hence, in this paper we aim to summarize the hints and the lessons learned from the organization of such events by providing as an example our experience in organizing this Doctoral Symposium (Ovaska, 2010).

The second concerns our feedback through this paper to the PhD students. In our opinion it is very useful for them to know the rules of the game in order to prepare and present a meaningful research plan, convince the reviewers and the attendees of the significance of their work, justify the results and value of the proposed work, and make a capturing/interesting presentation in front of the software domain experts. This will help students also in their future academic or industrial carriers.

1.2 Which Are the Lessons Learned?

Concerning the organization work, the co-chairs have learned that there is little material (except the Web sites of the conferences and workshops) on the organization rules of Doctoral Symposiums in general, and on the evaluation criteria in particular.

Concerning the feedback for the PhD students, the co-chairs have to outline that PhD students are often confused and have difficulties to present a research agenda for three-four years. Especially, describing the expected results and their validation seems particularly difficult for students. However, they are rather familiar with presenting workshop and conference papers which have a different objective. They also have difficulties to interact with domain experts (potential stakeholders) due to various reasons.

This paper summarizes the experience of organizing a Doctoral Symposium from the co-chairs point of view. The paper addresses aspects concerning both organization aspects and lessons to be learnt by the PhD students.

The rest of the paper is organized as follows. Section 2 presents the main steps of organizing a Doctoral Symposium, as well as the motivation behind these steps. Section 3 describes an actual example by focusing on the evaluation of the contributions. The lessons to be learnt are listed in Section 4. Conclusions are dealt with in Section 5.

2 ORGANIZATIONAL HINTS

This section lists the meaningful aspects the co-chairs have to address during the organization of the Doctoral Symposium.

2.1 Call for Contributions

The call for contributions should indicate the macro research topics of the Doctoral Symposium. Usually, they are the same or a sub-set of the topics of the main conference which the symposium is co-located with. This aspect is important because it is closely related to the expertise of the program committee members. Typically, part of the main conference program committee members also serve as the evaluators of PhD students research plans and are available and committed to face discussions with PhD students.

The call for contributions should mention the minimum and maximum length of the expected papers. The length may be imposed by the editors of the proceedings if the papers are accepted to be published in the conference proceedings or in a co-related publication. If no constraints are imposed, the contributions should be long enough to present the main aspects of the students' work, and in the same time as short as possible in order to avoid the insertion of details.

The call for contributions should clearly state what main aspects are to be addressed in the contribution and further used in the evaluation process. It is fair that the students know in advance the rules of the game. Based on their experience and after making a survey on the call for contributions for Doctoral Symposiums co-located in main software engineering conferences (e.g., ICSE (ICSE, 2013), FSE (FSE, 2013), ASE (ASE, 2013), ASWEC (ASWEC, 2013), APSEC (APSEC, 2013), WICSA (WICSA, 2013)) the co-chairs consider as significant the following ones:

- *the problem to be solved, its location and importance in the research field*; this shows the ability of the PhD students to focus on a problem, to locate it in the research field and to evaluate the possible impact if the results are achieved;
- *previous work*, which has addressed similar problems explaining why they have not been previously solved; this shows the ability of the PhD students to understand the research field and to avoid already known problems and mistakes;
- *the proposed approach*; this shows the ability of the PhD students to find appropriate solutions to problems and to make a medium-long term plan for achieving the identified solutions;
- *the expected results*; this shows the ability of the students to identify the impact and the value of their work, as well as the expected results after investing a significant effort to solve a non-trivial problem;
- *a plan for the evaluation of the results*; this shows the ability of the PhD students to sustain and demonstrate concretely the obtained results; this request is usually consciously or unconsciously avoided by the students, hence it has been decided to explicitly introduce it in the call for contributions.

Another significant aspect mentioned in the call for contribution is the stage of the PhD work. This influences the evaluation process (see Section 2.3).

2.2 Program Committee

The program committee should fulfil at least three

different criteria:

- members should be known experts in the field;
- the committee should be balanced and include members from academia and industry;
- members (or at least part of them) should be present during the workshop to discuss directly with the students; this is not easily achievable also because of the parallel events (e.g., sessions, workshops) usually organized during conferences, and hence experts may be involved in various events contemporaneously.

2.3 Evaluation Criteria

Based on the current stage of the PhD work, the five main points listed in the call for contributions are translated into the following evaluation criteria:

- problem addressed and its link to the Doctoral Symposium topics;
- motivation of the problem as an open research issue;
- importance of the problem in the research field;
- identification of the main related works;
- a proposed research plan with milestones for evaluation;
- appropriateness of the research plan for the problem;
- expected results;
- evaluation plan of the expected results;
- appropriateness of the evaluation plan for the results.

Furthermore, the clearness and the presentation quality are added to this list of evaluation criteria.

These evaluation criteria should be made available to the reviewers to achieve a homogeneous evaluation and as objective as possible results.

The co-chairs decided to not accept contributions which are in a very early stage (less than 6 months) because it is difficult for the students to present properly the last three points listed in the evaluation criteria. In early stage, students may start to know the previous work and try to identify a real problem and formulate it in a meaningful way. However, it is typical that they are not able to define a research plan, identify the possible results and/or make their evaluation plans. Furthermore, the co-chairs decided to not accept contributions which are in the advanced stage (less than 6 months to their finish) because the symposium date is usually a couple of months after the submission of the contributions and the feedback provided by the program committee members may hardly influence the thesis even if problems or open issues are identified.

The program committee members' feedback and the impact of the feedback on the students work are different in an early stage and in an advanced stage. In an early stage, the feedback may influence significantly the students' work starting from the definition of the problem until the evaluation plan of the results. In an advanced stage, the feedback may still influence the proposed approach and may focus on the results and their evaluation plan.

2.4 Letter of Recommendation

The role of the letter of recommendation is to confirm that the submitter is currently a PhD student. In addition, it should indicate the current stage of the PhD student's work and provide a qualitative and brief evaluation of the work done until the submission of the contribution.

2.5 Presentations and Feedback

The students' presentations during the PhD workshop should be perceived by the students as a short version of their final dissertation presentation. For example, a solution may be to allocate 45' to each student, 25' minutes for the presentation and 20 minutes for discussion. The discussion should consider both positive and negative aspects of the presented work in a constructive way.

2.6 Extended Abstract for the Doctoral Symposium

The co-chairs have to prepare an extended abstract about the Doctoral Symposium, abstract which is included in the conference proceedings. The extended abstract summarizes the call for contributions, introduces and thanks the members of the program committee for their collaboration, and provides an overview of the accepted papers (because the Doctoral Symposium papers are not always included in the main conference proceedings) (Ovaska, 2010).

3 AN ACTUAL EXAMPLE

In the case of this Doctoral Symposium, the received contributions were eleven, among which seven have been accepted for the presentation during the Doctoral Symposium. Three of accepted papers got an "accept" decision, while four of them a "weak accept" one.

The strong points of the "accepted" papers can be summarized as follows:

(1) real open problems identified,

(2) clear presentation of the problem and of the plan to address it,

(3) meaningful case study presentations,

(4) proper identification of the expected results,

(5) sustainable evaluation plan of the results.

However, also for the accepted papers improvement aspects have been suggested:

(1) extend the evaluation plan to further case studies or application domains (because some of them seemed narrow-scoped),

(2) discuss the impact of the results both from theoretical and practical points of view,

(3) delimit better the contribution of the PhD work if it is developed in the context of a research group/project which involves many researchers.

The strong points of the "weak accepted" papers can be summarized as follows:

(1) meaningful research question identified,

(2) potentially good solution plan introduced,

(3) acceptable expected results,

(4) promising evaluation of the expected results.

The limitations of the "weak accepted" papers concern:

(1) the unclearness of the overall presentation,

(2) poorly or very briefly addressed some of the required points explicitly indicated in the call for contributions,

(3) missing details for the comprehension of the overall approach,

(4) English-language presentation problems (which may contribute to the low quality related to the first three mentioned problems),

(5) format presentation problems (e.g., the requested format has not been adopted and hence, the length of the contribution is less or longer than the requested one).

The main reasons why four of the papers have been rejected are the following. One paper was not focused on the conference topics (e.g., the main stream of this software engineering conference research topics) and hence, on the topics of the co-located Doctoral Symposium. The second rejected paper was in a very early stage and the evaluation criteria were not met. In spite of the fact that the problem addressed and the idea behind the solution were very challenging, the paper was rejected. The third rejected paper did not describe clearly the addressed problem, the proposed approach, and the expected results. Its presentation quality was also quite poor. The fourth rejected paper was a particular case because it was a short version of the paper submitted and accepted to be presented during the main conference. In addition, it did not address all the requirements specified in the call for contributions, being thought from the beginning for the main conference.

The qualitative evaluation (considering as values: excellent, good, fair, and poor) of the contributions is shown in Figure 1.

Paper Number	Current Stage (Years)	Expected date of dissertation	Problem Addressed	Motivation of the Problem	Importance of the Problem	Related Work	Proposed Plan	Appropriatness of the Proposed Plan	Expected Results	Evaluation Plan	Appropriatness of the Evaluation Plan	Presentation Quality	Evaluation Reviewer 1	Evaluation Reviewer 2	Evaluation Reviewer 3	Additional Comments	Final Decision
1	2	2011	Execllent	Good	Excellent	Good	Execellent	Good	Good	Fair	Good	Good	Accept	Accept	Weak Accept		ACCEPT
2	1,5	2011	Good	Good	Good	Good	Good	Good	Good	Fair	Fair	Good	Weak Accept	Weak Accept	Weak Accept		ACCEPT
3	1	2012	Good	Good	Good	Good	Good	Good	Fair	Fair	Fair	Good	Weak Accept	Weak Accept	Weak Accept	The requested format has not been respected.	ACCEPT
4	0,5	2013	Good	Good	Good	Good	Fair		Fair	Fair	Fair	Good	Borderline	Weak Accept	Accept		WEAK ACCEPT
5	2	2011	Good	Good	Good	Good	Good	Good	Fair	Poor	Poor	Fair	Borderline	Accept	Weak Accept	Parts of the required points are mixed.	WEAK ACCEPT
6	1	2012	Good	Good	Good	Fair	Good	Good	Fair	Poor	Poor	Fair	Accept	Weak Accept	Borderline	Parts of the required points are mixed.	WEAK ACCEPT
7	1	2012	Good	Good	Fair	Fair	Good	Good	Fair	Fair	Fair	Fair	Borderline	Weak Accept	Weak Accept	More than a research topic it seems to me a case study.	WEAK ACCEPT
8	2	2011	Good	Good	Good	Fair	Poor	Poor	Poor	Poor	Poor	Poor	Weak Reject	Weak Accept	Borderline		WEAK REJECT
9	2	2012	Good	Good	Good	Good	Good	Fair	Good	Good	Fair	Fair	Weak Reject	Weak Reject	Weak Accept	The paper does not match the PhD workshop topics.	WEAK REJECT
10	0.5	2013	Good	Good	Good	Fair	Good	Fair	Poor	Poor	Poor	Good	Reject	Weak Reject	Borderline	Too early stage.	REJECT
11	2 years	2011	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Accept	Weak Accept	Weak Accept	An extended version of this paper has been presented and accepted to the main conference. Terminology inconsistencies.	REJECT

Figure 1: Example of evaluation results concerning the actual example.

4 LESSONS LEARNED

This section summarizes the main lessons the organizers and the PhD students should learn from and for a Doctoral Symposium.

4.1 Lessons to Be Learned by the PhD Students: Preparation of Contributions

PhD students interested in presenting their work in the context of a Doctoral Symposium should first understand the following hints:

1. Read carefully the call for contributions. First, meet the topics and then try to address all the requirements.
2. A Doctoral Symposium paper is not just another paper. PhD students should be really motivated and interested to receive a valuable feedback from experts to improve their three-four years work. It should describe a research plan/agenda, which may address one or more challenging research problems connected among them and which should be developed through three-four years. It is true, three-four years may seem a very long period especially in the IT world, and this may lead to changes in the agenda. However, the PhD students should be able to provide and sustain their research agenda also in a changing research environment.
3. The PhD research may be co-located within one or more research projects. An important observation should be made here: the projects should not be confused with the research plan/agenda of a PhD student. Students should clearly state their own agenda and the link with the co-related projects. In addition, a PhD student should clearly delimit his/her own contribution from the other members of the research group.
4. Do not try to invent problems. Try to find a real one, also considering the interaction with the industrial partners and lessons learned in industry settings.
5. The page limit may be a problem. Try to be concise, and in the same time precise.

4.2 Lessons to Be Learned by the PhD Students: Presentations and Feedback

The PhD students who have their work accepted to be presented at a Doctoral Symposium should consider the following hints:

1. Each presentation should indicate the potential application domains of the results of the research work and describe how the results will be validated in a case study or a set of case studies. Even if in the paper there is no space for such information, the presentation should include application domains and case studies.
2. PhD students should listen carefully to the questions and the comments made by the audience (program committee members and other experts) and only after the interlocutor has finished the question/feedback try to answer and provide further details. PhD students tend to answer immediately showing that they have quickly understood the intention of the interlocutor and that they have already thought about the raised problem. Students should admit that the audience may know more on a particular topic than themselves, even if there are persons in the audience who do not publish papers in this domain, but work in this domain.
3. PhD students should consider their interlocutors as friends, not enemies, which try to help them by even making uncomfortable questions and comments. Furthermore, they should establish closer contacts with (at least part of) the program committee members.
4. PhD students should also listen carefully to the other PhD students' presentations and ask questions about their work because they can learn about the more advanced PhD students' work and avoid repeating the similar problems in their own work. Furthermore, the presenters will be their future colleagues.

4.3 Lessons to Be Learned by the Doctoral Symposium Co-chairs

The organizers of a Doctoral Symposium should consider the following aspects:

1. Insert a request in the call for contributions which asks for a gantt concerning the main milestones for the proposed plan/agenda. Or, require the presentation of the gantt during the Doctoral Symposium.
2. The letter of recommendation has currently a formal role, indicating the submitter is currently a PhD student and that the research field is the one described in the paper. In addition, it specifies the year of the PhD research. This letter should have a more important role. It should

provide an overview of the student research and organizational abilities and a qualitative evaluation of the current status of the PhD student's work. It should also state the collaborators in this work (e.g., MsC students making their final thesis on topics concerning the research plan, as well as senior researchers in case the PhD work is co-located with an European or National project).

3. Require a list the accepted and submitted publications concerning the PhD research plan.
4. Provide a template for presenting the PhD research plans and clearly stipulate that it should be strictly followed.
5. Encourage PhD students' participation by allocating them one of the papers for evaluation. In this way, a PhD student can practice his/her competence on making analysis and representing the evaluation results in a critical, but polite manner and discuss about them with the program committee and workshop members.
5. Be always two or more co-chairs especially if the organizational period (e.g., review process) may include holidays and one of the co-chairs may not be available for several days.

5 CONCLUSIONS

PhD students are the talent of the future who shall solve the emerging challenges in increased complex situations in innovative ways. Thus, it is extremely important that they can get all the support for making their journey towards the core of their research community as fruitful and easy as possible. The primary intend of this paper is to support PhD students in making top-class research and to provide the guidelines for the chairs of Doctoral Symposiums to organize successful events. Indirectly, this kind of guidance can influence the quality and effectiveness of research, which in turn will have a positive impact on IT based innovation and business as well.

ACKNOWLEDGEMENTS

We would like to thank the conference organizers for giving us the opportunity to chair this event, and the program committee members for their valuable collaboration.

REFERENCES

Almi, N. E. A. M., Rahman, N. A., Purusothaman, D., Sulaiman, S., 2011. Software Engineering Education: The Gap Between Industry's Requirements and Graduates' Readiness. In I*EEE Symposium on Computers and Informatics*, pp. 542-547.

APSEC Series, 2013. Asia-Pacific Software Engineering Conference. http://apsec2013.eng.chula.ac.th/.

ASE Series, 2013. IEEE/ACM International Conference on Automated Software Engineering. http://ase2014.org/.

ASWEC Series, 2013. Australian Software Engineering Conference", http://www.aswec2014.org/.

ECSA Series, 2013. European Conference on Software Architectures. http://ecsa2014.cs.univie.ac.at/.

FSE Series, 2013. ACM SIGSOFT International Conference on Foundations of Software Engineering. http://fse22.gatech.edu/.

ICSE Series, 2013. International Conference on Software Engineering. http://2014.icse-conferences.org/.

Ovaska, E., Raibulet, C., 2010. Doctoral Symposium of the European Conference on Software Architecture 2010. In *Proceedings of the Fourth European Conference on Software Architecture: Companion Volume*, ACM International Conference Proceedings Series, pp. 1-3.

Subrahmanyam, G. V. B., 2009. A Dynamic Framework for Software Engineering Education Curriculum to Reduce the Gap between Software Organizations and Software Educational Institutes. In *Proceedings on Software Engineering Education and Training*, pp. 248-254.

Villavicencio, M., Abran, A., 2012. The Necessary Software Measurement knowledge in Software Engineering Education from the Practitioners' Point of View. In *Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 1-5.

WICSA Series, 2013. Working IEEE/IFIP Conference on Software Architecture. http://wicsa2014.org/.

Xueyun, J., Zihui, C., 2010. A New Way to Software Engineering Education. In *Proceedings of the International Conference on Educational and Information Technology*, pp. 1-3.

Adopting Commercially Inspired Practices Within an Academic Teaching Course

A Case Study of a Computer Games Engineering Degree

G. Ushaw, W. Blewitt and G. Morgan

School of Computing Science, Newcastle University, Newcastle, U.K.

{*gary.ushaw, w.f.blewitt, graham.morgan*}*@newcastle.ac.uk*

Keywords: Software Engineering, Project Based Learning, Engineering Education.

Abstract: A case study of a computer games engineering course is presented. The course has been designed with close input from industry and is achieving a high rate of success in the number of graduates being recruited by the target industry. The organisers of the course have extensive experience in both the software engineering industry, and in delivering academic teaching. These experiences are combined so that commercial software development practices, technologies and philosophies are adopted throughout the delivery of the academic course. The paper discusses the specifics of how and why this was achieved, and uses the Game Engineering course as an exemplar for encouraging the adoption of commercially inspired techniques within the teaching of software engineering and computer science more generally.

1 INTRODUCTION

Learning to program effectively is recognised as one of the most difficult tasks facing a prospective software engineer or computer scientist (**?**; **?**) (it has been claimed that a novice programmer may take ten years to become an expert (**?**)). Teachers must regularly update and redesign courses on this topic to make the learning process more accessible, and to maintain relevance to the latest developments in the field. Further to this, the software industry has high expectations of specific skills in graduate software engineers. Some of these expected skills are specifically technical, while others are related to good working and engineering practices.

Teaching of the specific technical aspects of programming is a challenging prospect, that can be researched, defined and updated to meet both academic and commercial needs. The aspects related to the practicalities of working within a larger software engineering environment are more difficult to define and present interesting challenges to teaching within an academic environment. It has been widely noted that there is unease from industry on the practicalities of what is taught on university courses (**?**; **?**), and a general acceptance that new graduates need varying amounts of specific training after joining a software development company. Anecdotally, academics can be seen as existing in their ivory towers with little interest in the practicalities of industry, while the commercial sector is at the coal-face with little interest in seeing what is beyond the most immediate challenges. Of course, these are wild exaggerations.

The authors of this paper have a combined background in both the software industry and academia, specifically in the field of computer game engineering. We propose that a more fruitful relationship between academia and industry can be fostered by adopting and promoting aspects of working methodologies and philosophies from industry within the academic course framework. This, coupled with the utilisation of technologies, equipment and software tools which are industry standard, should result in graduates who are better prepared for a career in software engineering, and are perceived as being more employable by their target industry.

In this paper we present a case study of a Computer Games Engineering Masters course which has been running for over five years, and enjoys high levels of support from the games development industry. This has been achieved by working closely with prominent members of the industry, and by combining commercial working practices and philosophy with the pedagogical demands of an advanced course on software engineering. The use of commercially inspired practices within the academic teaching has led to a success rate of over 70% of graduates gaining employment in the target industry within six months

of graduation (this compares to an average success rate of around 12% from game development courses across the UK (**?**)). The Games Engineering course is used as a case study, but the general points made are applicable to the wider fields of computer science and beyond.

2 METHOD

In designing the course in a way that incorporates commercially inspired practices, while maintaining pedagogical standards, a number of different elements were addressed. In this section, each of those elements is described in turn, focusing on the details of how and why each aspect was implemented.

2.1 Industrial Involvement

The most important way in which we have ensured our course incorporates commercially inspired ideas is to directly involve industry in the course. Industry involvement with the course is at the core, curriculum level, and not limited to mere endorsements. These industrialists are an integral part of the degree. From helping to identify course content to providing jobs for graduates, industry is involved at every stage of the course.

2.1.1 Industrial Advisory Board

The course is guided by an industrial advisory board. The board consists of experienced software engineers from industry who meet once a year to review course content. When inviting a developer to join the board, we focus on technical people (i.e. lead programmers and technical directors rather than managers or academic liaison), as the course content is heavily technical.

Commercial software developers are very busy people. When dealing with our board members, and especially when approaching a potential new member, a great deal of care is taken in respecting their crammed timetables and priorities. In the short term, the industrial advisers are providing a favour for the course; any benefit for them is much longer term (i.e. more employable graduates). Keeping commercial representatives from industry involved in the course is time consuming and requires significant effort, but the long term benefits for the course justify the effort required.

The presence of the industrial advisory board clearly states two things about the course. Firstly, it states to potential students that this is a course which is taken seriously by the target industry, increasing the chance of employment on graduation, and therefore encourages them to engage with the academic content. Secondly it states to members of that industry that the directors of the course take their advice and opinions seriously, crafting graduates in a manner which is beneficial to them.

2.1.2 Industrial Seminar Series

A series of seminars are presented by the members of the industrial advisory board to the student class (the seminar is optional for the industrialists, but most contribute). These seminars are typically on either a technical subject or an organisational one, and are designed to allow representatives of industry direct interaction with the student cohort. The speakers can use the seminars as a window to directly inform the students about an aspect of the subject which they feel is especially relevant.

Further to this, a session immediately after each seminar provides a more informal feedback opportunity between the speaker (often a potential employer) and the students. This session is used for further Q&A, and for the industry speaker to see the students' work and provide comment and feedback on it. One cannot overstate the impact that leading video game developers have on students. Video games is a creative industry and many developers are "famous". A faculty member indicating that a piece of work is good is nothing compared to an industry expert praising a student's work.

2.1.3 Industrial Placement

Credit for the second half of the Masters course is achieved from an individual project and dissertation. It is strongly encouraged that this project work takes place within a commercial software development studio if possible and practical. We have found that our industrial partners are very likely to take on one or more student to work on a project within their organisation. The combination of showing that we can evolve the course content in light of their input through the industrial advisory board, and the face-to-face interaction that we provide through the industrial seminar series, leads to an increased level of trust and interest in taking on our project students.

We decided early in constructing the course that any paperwork associated with an industrial placement should be kept to an absolute minimum for the industrial partner. This applies equally to any bureaucracy pertaining to the logistics of the student working away from the faculty, and to any assessment of the student's work from an academic perspective. In

our experience, reassuring a potential provider of a student placement that there is minimal red tape involved yields more interest in providing a placement. The student project must be subject to scrutiny and assessment for academic purposes, but this is achieved through direct dialogue with the industrial supervisor rather than supplying forms to be completed.

Building fruitful two-way relationships with industrial parties is not an easy task. The authors have many years of experience working within the target industry and, consequently, a wide network of contacts which can provide a rich source of collaboration. However these contacts are not taken for granted, and we are constantly aware that our relationships now, as academics, are different to what they were as industrialists. It should be apparent from this section that we take great care to treat industrialists with respect and to demonstrate that we understand their priorities, while ensuring the provision of an academically rigorous programme of study.

2.2 Hardware and Tools

A key feature of the case study course is that the hardware and software tools employed by the students during practical sessions are equivalent to those used by the industry. A further key feature is that the use of these tools is focussed on education rather than training - ii.e. the students are taught why the tool is useful and how it provides that utility, rather than a more straightforward guide to how to use a particular tool. The intention is that this approach results in graduates who can take on a new tool with confidence as they understand what it is for, and how it relates to the tools with which they are already familiar.

A simple example is the use of source control for all practical work. Any commercial software developer employs a source control system, and there are many professional packages available. From the first day of our course we issue each student with a source control account, and instruction on how and, more importantly, why to use it. This focus on the benefits of using source control, and the mechanism by which it is achieved, ensures that the choice of which professional source control package to use during the modules is almost irrelevant. When entering industry, the graduate will expect to use source control as part of the daily programming routine, and will understand and appreciate the reasons for doing so; the specifics of the particular solution will then be picked up within a few hours of use.

Combining the use of industry-standard tools with discussion of why they are useful, and how they work for the developer (rather than focussing on how the developer interacts with the specific tool) extends to all aspects of programming during our course. This includes the hardware itself (in our case games consoles, high-end PCs, tablets and smart-phones), and the development environment (compilers, debuggers, SDKs, IDEs, etc). The intention is that, by educating the student about why a tool is useful, and how it works, the resulting graduate will be much more comfortable moving onto new development environments or hardware, due to a better understanding of what the new technology is actually doing.

2.3 Collaboration Not Competition

Working within a commercial enterprise is a collaborative process, often involving work with developers elsewhere on the planet (**?**). In the authors' experience, the majority of students regard coursework and practical sessions as a competition with their classmates. In designing our case study M.Sc. course we have encouraged a collaborative atmosphere throughout. This applies to the nature and openness of the discussion and tutorial sessions, and the layout of the desks in the practical teaching environment (desks should face each other, rather than face a wall, to encourage student interaction). Each student has a workstation within the lab, which is where the entire day is spent, as would be the case in a commercial software development organisation. Lectures take place within this space (i.e. the teachers come to the students, rather than the students go to each lecturer's theatre). Students are encouraged to carry out all practical work in the lab, rather than taking it home. This approach is intended to imbue the class with a sense of belonging (i.e. it is *their* lab), and give the environment the feeling of a shared workplace.

It is important to note that the collaborative experience is not limited to the aspects of the course which include a team project. We encourage discussion and problem solving between students during individual coursework and practical sessions. While working within a commercial software development enterprise, one of the main tools available to a software engineer is the developer at the next desk: many problems can be resolved by a brief discussion with a colleague.

Team projects are a feature of most computer science degree courses, and are regarded as a crucial step toward becoming a fully-developed software engineer (**?**). While working in the software industry, the authors interviewed over one hundred graduate programmers from many educational institutions over a number of years. Almost without exception, when questioned about a team project, the response

involved a claim that the interviewee did almost all of the work, and the rest of the team did almost nothing (indeed, the author stopped asking about team projects in interviews as this repeated response rendered it pointless). By encouraging a collaborative atmosphere throughout the course, we hope to achieve a better group dynamic during the parts of the course involving teamwork, and a more positive experience of how teams should work together toward a common goal, which should then be apparent in subsequent job interviews.

Encouraging collaboration may also encourage plagiarism. Careful design of coursework, and clarity on which aspects gain marks, mitigates this issue. As discussed later in this paper, coursework is defined in a purposefully open-ended way so that each student can make decisions which test their own individual ability. The coursework does not have a right or wrong set of results, but consists of a framework in which to develop many related ideas. Consequently no two pieces of student coursework should be similar; if similarities are apparent then a code inspection and comparison should help resolve any potential plagiarism issues.

When the team project commences, the students are already familiar with source control. Using a shared code-base for the team project provides additional experience with standard working practices, while also inculcating effective teamwork and intra-team learning. A less gifted student may learn from changes made to the project by more gifted students. Conversely the more advanced student may realise that code generated for a team must be straightforward to interface with, and clearly structured and commented for other team members to understand and debug.

2.4 Increasing Autonomy for Students

Many studies show that encouraging student autonomy leads to greater engagement with the subject (**?**; **?**; **?**). Our case study course has been designed in such a way that the student cohort is given increased autonomy in the nature and direction of their study and coursework as the modules progress. Preparing a student for the workplace entails instilling a degree of self-reliance and the ability to move forward from one task to another without direct instruction. Note that encouraging both autonomy and collaboration is not contradictory; a well-rounded software engineer knows when other members of a team can be useful in progressing a problem or decision (i.e. the engineer can autonomously decide to collaborate).

Our case study course in Computer Game Engineering consists of three phases, each allowing the students more autonomy than the last. The first phase consists of sequences of lectures stripped daily across each week of the modules, combined with practical sessions and tutorials directly related to the lecture topics. These modules provide the basis of knowledge and ability hat is required across the remainder of the course. The tutorial and practical work is heavily prescribed, with direct instruction on how to achieve success (further exercises and suggestions are supplied, to push the more gifted or committed students). The modules are assessed by a combination of written closed-book examinations, and practical coursework. The coursework definitions become more open-ended as the modules in the first phase progress, allowing for more varied features to be added, and individual research to be rewarded, but the work required to achieve a pass mark is clearly prescribed.

The major element of the second phase of the course is the team project. The teams are allocated by the course supervisor so that each includes students with a full range of ability. The definition of the team project provides a specific set of goals, but is not specific on how they are attained; it also provides a high-level design for the product which is to be developed, but instructs the team to design the detail themselves. This approach encourages a combination of teamwork and autonomy. The course supervisor is available throughout the week on an informal basis, and reviews progress more formally (offering advice where required) on a weekly basis. The assessment of this phase involves no written examinations, but includes a demonstration of the finished product, technical reviews of the implementation, and evaluation of how the team interacted. There is also an individual research paper required during this phase - the subject is prescribed in the definition, and training is provided in researching a subject and presenting the work in the form of a conference paper.

The final term of the course consists of an individual project and dissertation for each student who has successfully progressed from the first semester. This work is self-led research over a period of five months. By this stage, the students are working completely autonomously, with a fairly informal review once a week with the course supervisor. The project topic is chosen from a list of possibilities provided by the course supervisor (many are generated by industrial or academic partners). Assessment is split evenly between implementation and dissertation.

In the authors' experience, this course structure successfully introduces, and increases, the autonomy of each student to the point where they can work independently and reliably, while having the confi-

dence and self-knowledge to ask for advice or discussion (whether from academic staff or fellow students) when it is useful. An interesting observation is that, while the individual project work could at be carried out elsewhere, almost all students not engaged in industrial placement work in the teaching laboratory, treating it as their workplace for eight hours a day, five days a week. This professional attitude impresses our industrial guests to the laboratory.

2.5 Focussed Module Content

A core philosophy in constructing and updating our case study course is to ensure that the subject matter of each module is focussed on the course's overall objectives and learning outcomes. The aim of our case study course is to educate potential software engineers on the core technologies utilised in developing high quality video games. It would be relatively easy to consider the various aspects of that topic, and to amalgamate some existing modules from other masters level courses into a new course with a little tailoring. However the focus of these modules would not be sufficiently relevant to be accepted by our industrial advisory board.

Consider an example: artificial intelligence is a topic to which many university modules are devoted, and is also a technology utilised in modern video games. However, our industrial contacts (and the authors' own industrial experience) would agree that most of the content of a more generalised artificial intelligence module is not relevant to developing video games, due to practicalities of either performance, memory requirement or reliability (**?**; **?**). An artificial intelligence module for a video games engineering course must be created bespoke with that application in mind. This approach to module construction applies to other aspects of the course and, the authors would argue, to any course which is to maintain good relationships with a target industry.

Although broad subject study is encouraged in laying foundations (e.g., learning to program is a key skill to become a video game programmer), when focusing on video games the subject material is constructed solely with video games in mind. For example, building a physics engine for a fast-paced first person shooter is not the same as building a physics simulation. Physics engines are primarily concerned with an illusion of reality for real-time game-play whereas simulation is concerned with an illusion of reality for off-line analysis. The products arising from two such programmes of study will be distinct; although related, their focus is different and the output (learning outcome) is not compatible.

When developing the course structure, we could have included some "easier" subjects and avoided the more challenging aspects of the core technologies used in video game engineering. This approach may well have attracted more students, and therefore increased the revenue for the faculty. However such an approach would not provide potential employers with sufficiently skilled graduates, which would lead to a poor relationship with industry, and ill-equipped graduates attempting to enter the careers market.

2.6 Coursework as Portfolio

We have mentioned coursework throughout this paper as it is a vital aspect of assessing students ability on an industry-led engineering course. The coursework is designed so as to have a dual purpose: firstly to provide practice and to measure student ability, and secondly to furnish the students with a set of demonstrations and experiences for discussion during a job selection process.

Potential employers in the engineering and programming sectors want to know whether a graduate applicant is capable of understanding and exploring a technical topic in depth. In the authors' experience a major part of a graduate's interview is focussed on the final year project. The projects we offer on our course are designed so as to involve in-depth exploration of a subject; further to this, as many as possible are designed in partnership with either an industrialist or an academic from another department. Consequently, projects usually have an interested party who is considered as the customer. This approach imparts the project with a sense of professionalism, with a target standard to be met to an agreed deadline.

The pieces of coursework attached to each module are also designed to form part of a portfolio, for use by the student at interview. Each coursework definition includes aspects which will appeal to the target industry (admittedly this is easier to define for games engineering than many topics), and students are encouraged to provide videos of their work online, so that potential employers can see their work and discuss it at interview. Feedback from industry has been very positive toward the quality of the coursework and its accessibility.

Table 1: The number of students known to have gained employment in target industries or academia within 6 months of graduation for each year of the course. (*)In the case of 2012-13 less than 6 months have elapsed since graduation.

Year	Total Graduates	Known Employed	Total Percentage	Lost Contact	Known Percentage
2008-09	11	8	72.7	3	100
2009-10	4	4	100	0	100
2010-11	20	14	70	6	100
2011-12	29	22	75.9	5	91.6
2012-13*	12	8	66.7	0	66.7
Total	76	56	73.7	14	90.3

3 RESULTS AND DISCUSSION

In this section we present the success rate of graduates from our case study course who enter full-time employment in the target industry, followed by some anecdotal feedback from representatives of that industry. We then discuss the successful elements of the course and consider possible improvements alongside some words of advice for course designers looking for inspiration from industrial practices.

3.1 Graduate Recruitment

Table **??** shows, for each year the course has run, the number of students that graduated from the course, the number of those students that are known to have achieved employment within the target industry or academia within six months of graduation, and that figure expressed as a percentage of the total number of graduating students. The total over the five years that the course has ran represents a percentage success rate of more than 70%. This compares very favourably to the UK national average from computer games development courses of around 12% (**?**).

The figures in the "known employed" column of Table **??** only account for the students with whom we have maintained contact, whereas the "total graduates" column includes all graduating students, so the actual success rate is likely to be even higher than stated. Indeed if we only consider students with whom we have maintained contact, then the percentage success rate is close to 100% (as seen in columns 4 and 5). Furthermore the 2012-13 cohort has only recently graduated (i.e. less than six months prior to the time of writing), so this figure is expected to rise somewhat. The course is at Masters level, and entrance requirements are high; consequently the intake numbers vary somewhat over the years. We maintain a high entrance requirement because the work on the course is difficult, and it would be unfair to accept the registration of a less able student who is very likely to fail.

We endeavour to remain in contact with students after graduation, usually through a professional networking website. There are three reasons for doing so. Firstly this allows us to keep track of the success rate from our course, as illustrated above. Secondly, software development companies regularly approach us looking for potential employees; by maintaining contact with graduates we can match suitable candidates even after they have left the university. Thirdly, and with much longer term effect, when our graduates succeed in the industry the hope is that they will become involved in the course, presenting seminars and eventually recruiting graduates.

3.2 Industrial Feedback

Feedback received from industry is overwhelmingly positive on the quality of the graduates from our case study course. The high percentage of graduates gaining employment is evidence that the industry is impressed with the results of our approach. Many of the members of the industrial advisory board take on one or more student project during the course, which then invariably leads to an offer of full time employment after graduation. Furthermore, those employers return the following year, often looking to increase the graduate intake from our course.

The authors make a point of engaging with commerce at all levels, from multinational software publishers to local SMEs and start-ups. As we have a long career in the target industry, we have many contacts who may be interested in getting involved in our course, or employing graduates. Maintaining a presence at industry events is important, but we elect to keep contact informal and friendly. We have also found that word-of-mouth between developers, especially SMEs and start-ups, is very important and positive. In the words of one CEO of a local SME, the authors are "helping keep local game development alive, due to the number of high calibre graduates pro-

duced".

3.3 Discussion

It should be apparent from this paper that the authors have designed a coherent course consisting of a series of inter-related modules which combine pedagogic demands with commercially inspired practices. Care has been taken throughout the course to ensure that the students' core skills are gradually built up, and that an understanding of why those skills are important pervades both the taught subject matter and the coursework and practical sessions.

Our use of projects involving incremental autonomy is intended to motivate students, and to build up high-level programming skills throughout the course. During each module on the course students are engaged in activities that complete a scaffold of the overall pedagogic program. It has been noted (**?**) that care should be taken that the scaffolds and activities needed to complete the exercises are arranged carefully in order to stress the relevant content, and also to help those students acquire programming skills effectively.

The authors have found that industrial involvement and feedback should not always be taken at face value. We are wary of crafting module elements toward specific needs of a particular industrialist. In particular, we are regularly asked to provide some training and knowledge of specific tools, engines, or SDKs (not coincidentally the technology which that industrialist is currently using for a project). We are open to integrating such tools into individual projects where they are relevant, but our module content must focus on the science and engineering which is the basis for those tools, rather than providing hands-on training with them.

Perhaps the key factor which we have employed in building this course is maintaining a good understanding of what the targeted industry expects of graduates, in terms of their longer term prospects. We focus on core technologies and underpinning theory, rather than specific implementations. This approach puts an onus on the teaching staff to not only remain *au fait* with developments in the field, but to recognise which developments stem from underlying conceptual ideas. The course is entirely delivered by instructors who are actively involved in research in the core technologies which drive the industry; without that research backbone, we believe the course would be significantly less successful.

4 CONCLUSIONS

We have discussed how commercially inspired practices and philosophies can be integrated into an academic teaching course, in a manner which supports the pedagogic needs of the teaching institution and student cohort. We have used our own M.Sc. course in Computer Game Engineering as a case study, and have attempted to present our experiences in a way which is applicable more generally to computer science degrees and other subjects.

We have achieved a very high success rate of students gaining employment within an industry which is notoriously difficult to enter. Coupled to this, we have built up a strong relationship with industry incorporating technically minded representatives into the planning and execution of the course. The authors' own experience has influenced the style of teaching, and has resulted in successfully inculcating a professional work ethic in the students, achieved through treating some aspects of the teaching lab in a comparable way to a commercial development studio. This is coupled with our philosophy of teaching the more difficult and challenging aspects of game engineering, and having a high entrance requirement for students.

Student satisfaction is measured by the University via a series of module feedback questionnaires. This feedback is anonymous and is optional. The scores attained by the course and the teaching staff are consistently high over the five years that the course has ran. Invariably the highest student feedback scores in the school of computing science are attributed to the teachers on this course. Perhaps more importantly, the informal feedback received from students after a few months in their first job is consistently positive in terms of how well they were prepared by our commercially inspired teaching methods.

REFERENCES

Baekkelund, C. (2006). Academic ai research and relations with the games industry. *AI Game Programming Wisdom*, 3:77–88.

Blewitt, W., Ushaw, G., and Morgan, G. (2013). Applicability of gpgpu computing to real-time ai solutions in games. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(3):265–275.

Boud, D. (1988). *Developing student autonomy in learning*. Psychology Press.

Corney, M., Teague, D., and Thomas, R. N. (2010). Engaging students in programming. In *Proceedings of the Twelfth Australasian Conference on Computing Education-Volume 103*, pages 63–72. Australian Computer Society, Inc.

Hilburn, T. B. and Humphrey, W. S. (2002). Teaching teamwork. *Software, IEEE*, 19(5):72–77.

Jiménez, C. and Villalobos, J. (2010). Learning/teaching a computer programming course. *Analysis of State-of-the-Art Solutions for Personalised Learning Support*, page 3.

Livingstone, I. and Hope, A. (2011). Next gen.: Transforming the uk into the worlds leading talent hub for the video games and visual effects industries. nesta.

Matthes, F., Neubert, C., Schulz, C., Lescher, C., Contreras, J., Laurini, R., Rumpler, B., Sol, D., and Warendorf, K. (2011). Teaching global software engineering and international project management. In *Third International Conference on Computer Supported Education*.

Perkmann, M., King, Z., and Pavelin, S. (2011). Engaging excellence? effects of faculty quality on university engagement with industry. *Research Policy*, 40(4):539–552.

Reeve, J., Bolt, E., and Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology*, 91(3):537.

Reeve, J., Jang, H., Carrell, D., Jeon, S., and Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and emotion*, 28(2):147–169.

Vega, C., Jiménez, C., and Villalobos, J. (2012). Implementing an incremental project-based learning solution for cs1/cs2 courses. In *Second International Conference on Computer Supported Education*, pages 15–27.

Winslow, L. E. (1996). Programming pedagogya psychological overview. *ACM SIGCSE Bulletin*, 28(3):17–22.

Continuous Assessment in Civil Engineering Education
Yes, but with Some Conditions

F. J. Martín-Carrasco, A. Granados, D. Santillán and L. Mediero
Department of Hydraulic and Energy Engineering, Universidad Politécnica de Madrid (UPM), Madrid, Spain
{f.martin, a.granados, david.santillan and luis.mediero}@upm.es

Keywords: Continuous Assessment, Computer Tests, Final Grades, Higher Education, Technological Subject, Students' Learning.

Abstract: Adapting Spanish curricula to the European Higher Education Area (EHEA) implies the introduction of continuous assessment. Continuous assessment is generally considered to enhance students' learning. The new methodology contrasts with the traditional Spanish method of assessment, based only on exams. This paper compares the student's learning under these assessment methods in Civil Engineering (Civ.Eng.). The results of 16 consecutive years of assessment of a technological subject (Hydraulic Engineering) have been analysed. Assessment during the first 8 years was performed only by final exams; and onwards by classroom exercises, computer tests in a Virtual Learning Environment (VLE) and final exams. Rates of students that passed are clearly higher in this latter period, what seems to support that a better learning is achieved under continuous assessment. However, when analyzed in detail, it is found that exams scores are significantly lower during the continuous evaluation period. Sometimes the appearance of a higher grade may mask a lower level of learning. The reasons lie in the psychology of the students, since they reduce their effort once a satisfactory score is achieved in the periodical training. The paper examines what elements should be incorporated to continuous assessment to improve student's learning.

1 INTRODUCTION

In the convergence of the Spanish curricula to the European Higher Education Area (EHEA), the assessment methodologies have entered in a new framework, in which the development of skills and assimilation of contents is assessed continuously, giving feedback to students during the learning period. This change in the assessment methods has been fostered by the development and spread of new tools as the Virtual Learning Environments (VLE), which opened many possibilities for transmitting information and interacting with the students.

In this framework, continuous assessment has been adopted by Spanish universities, which assume that assessing students throughout the course with periodic tests and courseworks will enhance their assimilation of knowledge and development of skills (Delgado, 2005; Hernández, 2012).

Civil Engineering is imbibed in this change period. Current Civil Engineering Degree, which nowadays concedes the capacity to fully develop the professional activity, is being substituted by a Bacherlor's Degree (B.Sc., 4 years long, that conceded limited professional qualifications) plus a Master of Science (M.Sc., 2 years long) (MCIN, 2009a; MCIN, 2009b). In most Spanish universities, the B.Sc. is being implemented, while the M.Sc. is under preparation. Besides, the current Civil Engineering Degree is being in extinction process, which is done coordinately with the introduction of the new titles.

The implementation of the new M.Sc. is the next step towards harmonizing the titles to the EHEA. For instance, at the Universidad Politécnica de Madrid (UPM), it will be introduced the next academic year 2014-2015. Therefore is necessary to design the subject's assessment methodology. Within the abovementioned extinction process, some technological subjects of the Civil Engineering Degree have been applying a continuous assessment methodology on recent years, in order to better adapt to the new EHEA requirements. The analysis of this experience may be useful for identifying weaknesses of the continuous assessment when applied to technological subjects, which is the objective of this work.

Continuous assessment provides certain

advantages over other methods (García-Beltrán, 2002; Trotter, 2006; Isaksson, 2007; Joughin, 2009). The method enables the students to gradually assimilate knowledge and progressively develop their skills. Furthermore, the method provides information about the learning process. Students and teachers may be interested in this information which provides a feedback on the learning process. Moreover, the method is a student-centered learning system which offers a better preparation towards the final exam, as it may be similar to the exercises solved by the students in the periodic tests and courseworks.

Contrary to the previous advantages, several drawbacks have been reported (Yorke, 2003; Martínez, 2008), such as the difficulties to develop an ambitious theoretical content in the course program, since theory is partially substituted by practice. Moreover, continuous assessment cannot be effectively implemented in large groups, because individualized attention to the learners requires a substantial time commitment from academics.

Since a link between what the students must learn and what is actually being evaluated must be established, teachers must choose the most adequate assessment method. The assessment should be understood as a tool for improving students learning. (Dochy, 1997). To design a continuous assessment system, various evaluative activities with different relative weights must be defined (Gallardo, 2010; Gallardo, 2011). In addition to this, it has to be decided if the final grade is obtained only by the continuous assessment results, or if a final exam is also performed (Haghnegahdar, 2013). The use of this periodical assessment is not a substitute for final exams but is an important complement in order to define a students' learning focused assessment method (Dochy, 2007).

The M.Sc. Civil Engineering has a highly technological content. Students are required to develop the highest cognitive categories, such as: application, analysis, synthesis and evaluation (Bloom, 1956). Technological subjects are oriented to produce final designs or solve specific management problems. Those designs and problems are always unique and have to match with the site or system specific characteristics. In this type of subjects a final exam is highly recommendable. In this type of subjects, partial skills have to be joined to achieve a final objective. Different partial skills can be evaluated separately through continuous assessment, but their aggregation towards the final objective can be only evaluated in a final exam.

This paper contrasts the students' learning under two different assessment methods in a technological subject in higher education. In the first method the grade depends only on the final exam mark, whereas in the second, the grade depends on the combination of the final exam mark and the periodical exercises ones, as continuous assessment practices were implemented. The comparative analysis of the assessment results under both methodologies (continuous and traditional only by exam) may be useful for detecting critical points, which would help in designing and applying the assessment methodology on the future M.Sc. Civil Engineering subjects.

2 STUDY ON THE ASSESSMENT OF A TECHNOLOGICAL SUBJECT

Results of assessing a compulsory subject of the Civil Engineering Degree, Hydraulic Engineering, are described and analyzed. In line with the objectives, this section is organized as follows: first a brief description of the subject is presented, second the evolution of the assessment methods is described, third the factors which may influence the assessment are analyzed, fourth the students' performance results are presented and finally, the investigation leads to an interesting discussion regarding students' learning performance.

2.1 Subject Description

Hydraulic Engineering is a technological subject of the 5th year of the Civil Engineering Degree at the UPM. This degree consists of six academic years and a final project (syllabus is shown in Figure 1). The subject has a highly technological content. Students need to have a strong scientific basis which is provided during the first two academic years and a technical previous background that is acquired during the 3rd and 4th years of the degree. The bases of the syllabus of this degree are established by law (MEC, 1983). The basic academic data of the subject are:

- Compulsory subject of the Civil Engineering Degree.
- Teaching is annual and is divided into 2 terms.
- The form of teaching is 120 hours of lectures during 30 weeks and about 500 hours of student study (equivalent to 19 ECTS).
- The first term (15 weeks) covers the topic of dams.
- The second term (15 weeks) is dedicated to water

distribution systems, pumping technologies, hydropower and irrigation.
- The subject provides students skills and qualifications for: planning, developing, projecting, managing the construction and operating hydraulic engineering systems.
- The number of students for the 16 years of the analysis is, on average, 450.

SCIENTIFIC TRAINNING							
1ST YEAR	Algebra	Calculus	Physics	Chemistry	Technical drawing (I)		
2ND YEAR	Mathematical Analysis	Mathematical Methods	Mechanics	Building Materials	Technical drawing (II)		
TECHNICAL TRAINNING							
3RD YEAR	Statistics	Differential Equations & Numerical Methods	Strength of Materials	Geology	Electricity & Electronics	Topography	English (I)
4TH YEAR	Structures	Concrete structures	Hydraulics	Urbanism	Geotechnics	Economy	English (II)
TECHNOLOGICAL TRAINNING							
5TH YEAR	Steel Structures	Hydraulic Engineering	Roads & Highways Engineering	Port & Coastal Engineering	Transport Engineering	Art & History of Civil Eng.	Specialization in: - Transport - Structures - Hydraulics - Urbanism
6TH YEAR	Construction	Projects	Railroads Engineering	Water & Wastewater Engineering	Business Admin.	Laws	

Figure 1: Syllabus of the Civil Engineering Degree (MEC, 1983).

2.2 Evolution of the Assessment Methods

Before the EHEA, students were generally assessed only by a final exam. The final grades were proportional to the score on the exam. There were usually three final exams. The first exam (ordinary exam) was at the end of the second term in May or June. Students who failed this test had two resit exams, one in September and another in December. Marks were within the range 0 and 10. A 0 grade involves a very poor performance and 10 an excellent one. The pass mark is 5.

As a result of the adaptation to the EHEA, the continuous assessment method has been included in the evaluation process. The method considers several activities carried out by the students during the course, such as attendance to lectures, classroom tests or courseworks, and an ordinary final exam. The final grade is the weighted sum of the mark on the course activities and the score on the final exam. Students who achieve a grade lower than 5 fail the subject. In such case, learners have the two resit exams aforesaid in the previous paragraph and the final grade is equal to the mark on the exam.

The assessment of students for the 16 years of analysis may be divided into three periods:
- From the academic year 1994-1995 to 2001-2002, students were examined only by a final exam.
- A transition to continuous assessment started in the 2002-2003 academic year. The new methodology was completely implemented after two years.
- A continuous assessment period of 6 academic years, from 2004-2005 to 2009-2010. In this period, up to 40 monitoring exercises were proposed to be solved by the learners. Two types of exercises were performed: classroom tests and computer tests. Classroom tests were randomly proposed and solved by student during the lectures. The tests consisted of a theoretical or practical question related to the concepts that were explained during the lecture or in the immediately previous one. The exercises were completed in about 10 minutes. Computer tests were solved by students at home in a VLE (the Moodle application). The tests were composed of a unique exercise or several questions, both related to the lectures of the previous month. The exercises were available for one week to the students. They have to answer the proposed exercises in about one hour.

2.3 Analysis of the Factors Which Influence the Assessment

Several factors may influence assessment; among them are: those related with people involved in the educational process (students and academics), those related with contents and skills to be learned and developed, and those related with methodology (exams, exercises, etc.).

This section analyzes the progression of these factors during the study period. This analysis is necessary for discussing the results and drawing conclusions about the effectiveness of continuous assessment. The different facts evolved as indicated below:
- Students' profile and requirements to register in the subject have not varied. Pre-university profile has remained constant since 1996, when the later high-school education law became in force. In a similar way, the university profile has remained constant as well; there were no changes in syllabus neither in requirements for registering in the subject (four years of previous scientific and technical training).
- As stated above the syllabus of the Civil Engineering Degree didn't change during the period of study. That is to say, subject's contents and structure have not been altered. However, the contents were progressively adapted to the state of the art, but always following the main subject's topics (dams, water distribution networks, etc.). Total time devoted to each topic didn't change significantly.
- Additionally professional regulation has not been

modified during this period. Thus, subject's learning objectives have not change. In essence, students have had to acquire the same knowledge and skills.

- Academic staff has not suffered any important modification. The main academics have been in the same post during the period of study, and new staff has been recruited with the same profile.
- Assessment method shifted from a traditional, only by final exam, methodology to a continuous assessment methodology.
- Exams have kept its structure. The number of theoretical questions and practical exercises, the weights and the duration has been the same during the period of study. The difficulty of the exams has not increased. Proposed questions and exercises are used as base for preparing new exercises. New exercises were similar to the previous ones, since both look for accomplishing the same goals, because knowledge and skills to be assessed were the same.

2.4 Students' Performance Results

The results of these 16 years of assessment are shown in the next three figures. The rates of students that passed the subject during the period of analysis are displayed in Figure 2. This graph gives a first idea of the continuous assessment effectiveness. This first result has to be analysed in depth, to understand if the method accomplished its objectives. Further analysis is done with the help of other graphs, showing the evolution of the subject's grade and the evolution of the final exam's grade. These facts are summarized in Figures 3 and 4.

Figure 2 presents the rate and distribution of students that passed the subject during the regular academic period (ordinary final exam) and the students that passed after the extraordinary resit exams. In can be observed that once continuous assessment has been completely implemented (course 2002-2003) the rate of students that passed at the end of the regular academic period grown by 20%. It can be seen also that the rate of students that passed after the extraordinary resit exams is very small, ranging between the 2% and 4%. These data lead to the idea of the success of the continuous assessment method; but further analysis has to be done to estimate the accomplishment of the objective, which is not just a higher rate of students that passed but a better students' learning.

Figure 3 shows the evolution of the overall grades on the subject, discriminating the students that passed after the regular academic period and

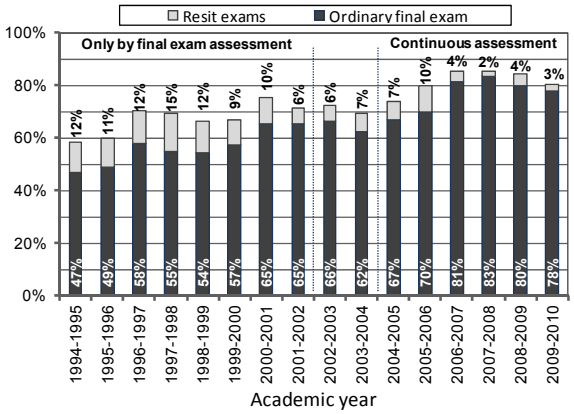


Figure 2: Rate and distribution of students that passed.

those that did it after the extraordinary resit exams. No upward tendency could be seen in the charts, with both lines remaining constant along time; which indicates that the assessment method does not influence the final grades of the students. It can be also seen that the students that passed after the ordinary call (average grade of 6.93) have a better performance than the ones that passed after the resit calls (average grade of 6.11). That was predictable, since the better students usually passed in the first call, making the difference between the average grades of the two groups.

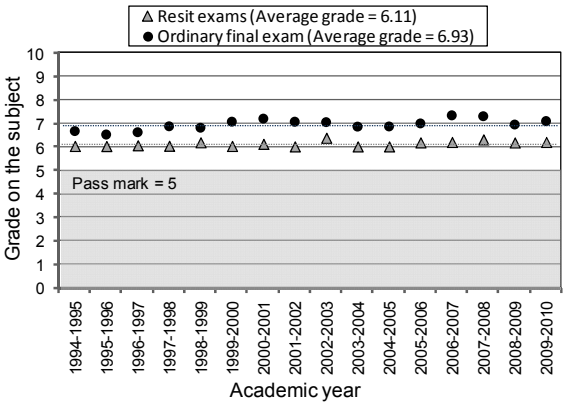


Figure 3: Evolution of the subject's grade.

Figure 4 is useful to understand in depth, the students' performance. It shows the evolution of the overall subject's grades compared to the evolution of the final exam's grades. Data of this figure has been computed only for the students that passed after the ordinary period. For the traditional assessment period both grades coincide, due to the fact that the only assessment tool was the exam. Then, is interesting to observe how a new tendency appeared when new continuous assessment tools begun to be

operative (2002-2003 and onwards). This new tendency is clearer when the new assessment methodology was totally implemented. As can be observed, the students have a poorer performance in the final exam during the later years of the study period, those which corresponds the continuous assessment methodology. For the last 6 years of the study period the average grade (6.93) is 1.61 point over the average grade on the final exam (5.32).

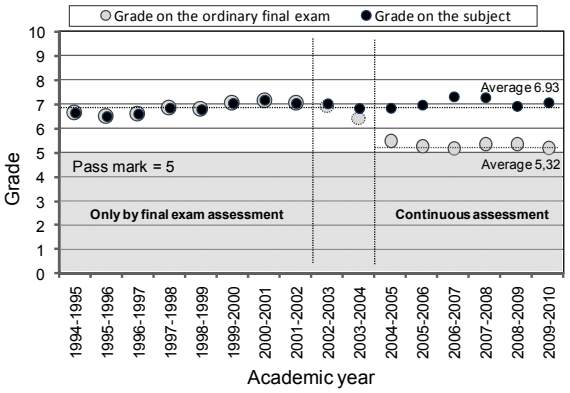


Figure 4: Evolution of the subject's grade vs. final exam's grade.

This fall of almost 25% in the final exam grades during the last 6 years, may be due to the change in the assessment model, because the other factors that may influence the learning process: the students (number and background), the subject (length, objective and contents) and the academics (main professors and background of the new ones); have not had significant changes during the whole period of analysis.

2.5 Discussion and Improvement Measures

It is acknowledged that the study have some limitations which should be consider when discussing the results. It is true that the long period of the experiment let to have more results on students' performance under the two methodologies; however it should be noted that during these long period the facts impact of the educational process may slightly change, impacting on results. Although the factors of the analyzed subject have remained stable through years, further analysis on small changes may be carry out to complement the results. The study has been done at an aggregate scale, comparing overall subject's grades and final exam's grades. Additional disaggregation of the analysis, regarding the evolution of the different questions and exercises, would be useful for detecting critical points on the students' learning process. Nonetheless, the experience on the assessment of a highly technological subject under these two systems and the comparison of their results provide relevant information; which may be useful for educators in order to improve the assessment methods of such type of subjects.

The analysis of the students' performance results leads to three interesting conclusions, which seem to be directly influenced by the assessment model:

- The number of students that passed the subject under the continuous assessment method has grown by 20% in comparison with the traditional assessment.
- The average grade on the subject throughout the years of study has remained steady (around 6.93 for the students that passed during the regular academic period).
- The average grade on the final exam under the continuous assessment has decreased in almost 25%.

In turn, aggregating these three conclusions it can be deduced that: although continuous assessment leads to better rates of students that passed, the performance of those students on the final exam is lower; so the initial objective of improving students' learning is not completely accomplished. There are three causes that explain those results:

- The first cause is located in the character of the subject. The objective of a technological subject in higher education is to solve a general problem or to produce a final design. This objective requires combining all the knowledge that is explained during the course. On another note, continuous assessment focuses on the day by day students' performance and may lose the view of the subject as a whole. This fact should be taken into account when designing the continuous assessment method. Thus, for technological subjects, midterm control exams may be introduced. These exams should require students to link together the contents that were previously explained, towards the resolution of a problem. So that they do not lose the overview of the subject.
- The second is located on the size of the group that is been assessed. In continuous assessment the students' learning process benefits from the feedback that is periodically given. The communication and the relationship students-academics are more difficult in large groups, so the effectiveness of this continuous assessment is reduced. In this sense, VLEs are useful tools that

simplify the work of academics and facilitate rapid feedback to students.

- The third one is located on the students' psychology. The periodic tasks that were performed during the course (room and VLE test), are assessed and provide a mark that is taken into account in the computation of the final grade. In this context students reduce their work towards the final exam, because they know that they could pass the subject with a lower grade in the final exam. So they relax and calibrate their effort. This circumstance should be considered when defining the tasks, the weights and the subject's regulations. For technological subjects, in which a final exam should be carried out, it is highly recommended to require an independent minimum grade in this exam.

3 CONCLUSIONS

During the last decade, the continuous assessment methods have been imposed in the new educational systems (as the EHEA). Presumably, under this assessment system, students are more involved during the course, resulting in higher rates of students who pass and higher grades. These factors are associated to a wider development of skills and a better learning performance. There is no doubt about the benefits of continuous assessment; but in technological subjects in higher education it should be implemented under some conditions. Otherwise it may lead to a poorer learning.

The case study shown in this paper proves that lower learning performances could be masked under higher rates of students that passed or higher grades. The reasons lie in the weighting of the different tasks and in the psychology of the students, since they reduce their effort once they achieved a satisfactory mark in the continuous training, loosing interest for the final exam. This final exam is an important assessment tool in technological subjects, in which students should link together all the parts of the course to solve a general problem or to produce a complete design.

To avoid this problem, some additional conditions should be introduced when designing continuous assessment, such as: to set midterm exams for providing an overview of the subject, to take advantage of the VLEs to offer a effective feedback to the students, and to set a minimum mark on the final exam, independent of the other course task' marks.

REFERENCES

Bloom, B. S., Taxonomy of educational objectives. Handbook I: *cognitive domains*. David McKay Co. Inc., New York, US.

Delgado, A. M. (Coord.), 2005. Competencias y diseño de la evaluación continua y final en el espacio superior. *Dirección General de Universidades*, Madrid, Spain.

Dochy, F., and Mc Dowell, L., 1997. Assessment as a tool for learning. *Studies in Educational evaluation, 23*, 279-298.

Dochy F., Segers, M., Gijbels, D. and Struyven K., 2007. Assessment engineering: breaking down barriers between teaching and learning, and assessment. In Rethinking assessment in higher education: *learning for the longer term.* Routledge, Oxon, UK.

Gallardo, E., Montolio, D., and Camós, M., 2010. The European Higher Education Area at work: Lights and shadows defining continuous assessment. *Revista d'Innovació Docent Universitària*, 2, 10-22.

Gallardo, E., 2011. ¿Existe relación entre la evaluación continua y los resultados de los alumnos?. *Revista Electrónica sobre la Enseñanza de la Economía Pública*, 8, 63-69.

García-Beltrán, A. and Martínez, R., 2002. The role of self-assessment in AulaWeb e-learning system. In European Distance Education Network (EDEN) *Annual Conference*, Granada, Spain.

Haghnegahdar, A., 2013. Alternatives to heavily-weighted final exams in engineering courses. *Teaching Innovation Projects, 3*(1), article 2.

Hernández, R., 2012. Does continuous assessment in higher education support student learning?. *Higher Education, 64(4)*, 489-502.

Isaksson, S., 2007. Assess as you go: The effect of continuous assessment on student learning during a short course in archaeology. Assessment and Evaluation in Higher Education, 33(1), 1-7.

Joughin, G., 2009. Assessment, learning and judgment in higher education: A critical review. In G. Joughin (Ed.), *Assessment, learning and judgment in higher education*, 13-27.

Martínez, M., 2008. La importancia de los nuevos modos de evaluación en el EEES. Una aproximación a las ventajas del uso del portafolio. *Revista de Enseñanza Universitaria*, 31, 62-72.

MCIN (Ministerio de Ciencia e Innovación), 2009a. Orden CIN/307/2009 por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de la profesión de Ingeniero Técnico de Obras Públicas. *Boletín Oficial del Estado*, de 9 febrero.

MCIN (Ministerio de Ciencia e Innovación), 2009b. Orden CIN/309/2009 por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de la profesión de Ingeniero de Caminos, Canales y Puertos. *Boletín Oficial del Estado*, de 9 febrero.

MEC (Ministerio de Educación y Ciencia), 1983. Orden por la que se aprueba el nuevo Plan de Estudios de la

Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos de la Universidad Politécnica de Madrid. *Boletín Oficial del Estado*, de 9 septiembre.

Trotter, E., 2006. Student perceptions of continuous summative assessment. *Assessment and Evaluation in Higher Education*, 31(5), 505-521.

Yorke, M., 2003. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501.

Game-based Learning in IT Service Transition
The Case of a Mobile Sales Service by a Small Team in Brazil

Thiago Paiva Brito[1], Josias Paes[2] and J. Antão B. Moura[3]
[1]*Graduate Course on IT Management, UNIPÊ, João Pessoa, PB, Brazil*
[2]*Unipêtech, UNIPÊ, João Pessoa, PB, Brazil*
[3]*Systems and Computing Department, Federal University of Campina Grande (UFCG), Campina Grande, PB, Brazil*
tpborion@gmail.com, josias.paes@unipe.br, antao@dsc.ufcg.edu.br

Keywords: Game-based Learning, BPM, ITIL, Software Deployment, Motivation, Deployment Time, Learning Curve.

Abstract: IT Service Transition (ITST) is naturally challenging because it usually involves changes that run counter customer and provider staffs' preconceptions, habits and established practices. Changes that affect implicit or explicit business processes (BP) adopted by the provider's Service Transition team are particularly daunting since they may impact the team's morale negatively and contaminate customer's personnel who might be anxious to start with. Inability to properly implement and manage changes due to Service Transition process adjustments and retooling may lengthen deployment time, lower quality and even cause the provider to fail. In order to efficiently handle such changes, the provider's ITST team must be motivated towards, trained in and quickly made proficient with new work tools, routines and practices. This paper provides preliminary evidence that blending Business Process Management (BPM) to gamification concepts and tools can accelerate learning in an IT Service Transition context. For that, we consider the case of a small IT Service provider in Brazil when transitioning a sales support IT service from Palm OS to Android devices.

1 INTRODUCTION

According to the IT Infrastructure Library (ITIL), the mission of IT Service Transition (ITST) is to deliver services that are required by the business into operational use (Macfarlane, 2007). However, due to the particularities of the target environment where each service will operate, Service Transition is likely to be the most critical phase of the IT service life cycle. ITIL documentation goes on to prescribe what needs to be done in terms of recommended Service Transition practices, but little is offered on how to, in fact, implement such practices. Nor could it be otherwise since implementation is to be subject to the provider's idiosyncrasies and preferences, service characteristics and clients (business) needs. Hence, the provider must design and adjust its processes to meet the Service Transition tasks and goals at hand. Further, making the IT staff learn (and adopt) new working routines is usually met by human resistance which in turn, increases risks of one making unsuccessful ITST efforts. Often and understandably, the designer ends up treating the results as trade secret. Thus, one is left with the option of experimenting with ITST Business Process Management (BPM).

We suggest that blending gamification to BPM may accelerate the learning curve for new Service Transition tasks and thus ease transition complexity and risks. Here, "gamification" as in (Werbach and Hunter, 2012; Deterding, et al., n.d.), is the use of game mechanics and game design techniques in non-game contexts (e.g., in an ITIL training context). The suggestion presupposes that motivational and training sessions for the Service Transition team through a gamified approach could trigger just the level of engagement one needs for the success of learning how to work with new transition Business Processes (BP) and tools. A gamified approach to ITST could create a little bit of fun and amusement around working processes (Werbach and Hunter, 2012), therefore, should help to break the ice and diminish the resistance to needed, progressive, and beneficial change to Service Transition processes management. The objective of this paper is to present and evaluate a game-based "learn by doing" approach for ITST.

Some works in the literature use games to provide an engaging, self-reinforcing context in which to motivate and educate players (serious games) (Kankaanranta and Neittaanmki, 2008). Other works simply try to engage users into work through fun (Castellani, et al., 2013). The bibliography on gamified education indicates that gamification aspects facilitate learning and working in professional and business environment in general (please see for instance, the proceedings of CSEDU 2012, 2013 – www.csedu.org). Not much has been published however, on applying gamification to ITSM contexts.

The game-based learning approach proposed here focuses on using fun as a tool for enabling changes in ITIL processes (ITST) from a possibly ad hoc, undocumented state to a properly managed process where performance can be measured, the actors involved get feedback about their work and company stakeholders perceive real gains. As such, it is hoped that the game-based learning method to be detailed here also contributes to the state-of-the-practice of ITSM (ITIL in particular).

Thus, the research question this paper addresses, in order to provide (preliminary) evidence that gamified education is also worthwhile in ITST contexts is: Can one provide evidence that gamification also makes learning and working more effective and efficient in the technical and specialized setting of IT Service Transition?

Here, we assume that preliminary evidence may consists of (1) signs of increased motivation by Service Transition staff for learning and applying new transition processes ("learn by doing"); (2) declarations by Service Transition team and company stakeholders that they are satisfied the game-based training approach helps them address complexity and risk issues successfully in a cost-effective manner; and (3) the approach renders verifiable results for the efficiency and effectiveness of the (newly adopted) ITST processes.

2 PROPOSED ITST BPM GAME

The proposed game-based approach to ITST is detailed here to the extent of supporting the evaluation of how the approach will: i) motivate and prepare a small ITST team to comply with new processes and tools and perform some less creative and repetitive activities with the required level of attention and care; and, ii) help to diminish team's resistance to improvements in the transition process provided by the BPM cycle (Rosemann, 2010). In short, the overall objective of the gamified approach is to ease learning and usage of ITIL's Service Transition processes.

Borrowing from game design, rather than using game-based technology, or other game related practices, and using some game elements (rather than full-fledged games) (Deterding, et al., 2013), the ITST gamified approach is designed and applied according to a somewhat empirical methodology and uses Points, Badges and Leader boards (PBLs) derived from a progression map (modeled as a BPM Notation diagram).

The learning approach calls for IT staff to initially study a BP diagram (BPD) which serves as a (process) progression map for navigation through the (ITST) game. To reduce resistance in using the "map", two basic prerequisites are to be met:

i. The diagram should provide a clear and unambiguous vision of the whole ITST process. Process documentation should also clarify the objectives and procedures for each activity the team is to perform.
ii. Improvements to any process should be communicated in a simple way since changes will be clearly shown in the progression map.

The reward scheme of the game-based ITST learning approach awards points as follows:

- For transition tasks concluded on time and which are of:
  - Low complexity: 5 points
  - Medium complexity: 10 points
  - High complexity: 20 points
- Bonus:
  - Tasks concluded before deadline: +5 points
- Unlimited additional points will be given for exceptional performance in any task (at the discretion of the team's manager or other higher ranking company officer):
  - For every customization process bug found and corrected: +10 points
  - For every automated test coded for a specific client customization: +10 points
  - For every automated test coded for bugs in the base system: +20 points
  - For every suggestion of improvement in the whole transition process: +20 points

The amount of accumulated points will reward the "player" with "Ninja Badges" as indicated:

- 150 points: Badge "Ninja Apprentice" (white belt).
- 450 points: Badge "Ninja Novice" (yellow belt).
- 750 points: Badge "Ninja Deputy" (green belt).
- 1500 points: Badge "NINJA!" (black belt).

The design assumes that the amount of points

accumulated by a BP actor will serve as an immediate feedback about her/his learning and work performance at task completion time, functioning as extrinsic motivation by external regulation (Deci & Ryan, s.d.).

An additional assumption is that extrinsic motivation also serves as "identification of merit" (through awards of points and badges). A given badge identifies a certain professional skill level on IT service transition. Identification (Deci & Ryan, s.d.) through merit points and badges could boost ITST team members' sense of security about their professional ITSM skills.

With a badge structure assigned at each punctuation level and classifying the ITST BP actor as a "Transition Ninja", the game-base ITST learning (and practice) approach acts on the extrinsic motivation through introjection. Here, we expect that a future recognition as "Ninja" will be introjected by pride and need for professional evolution.

The intrinsic motivation of the team will be stimulated by unexpected bonus awards for each activity completed with exceptional performance and other natural stimuli coming from their perception of fun in playing the (ITST BP) learning game.

2.1 Implementation

Implementation of the proposed gamified ITST learning approach yielded a PBL tool that registers and maintains the scores of each player (ITST BP actor). For the case study to be examined next, the implementation efforts consisted of building a shell of game elements (to support the proposed game mechanics) around a BPM tool. Basically, three new screens were added to a Python/Django system to allow for (see also Figures 1): i) registration of a new player; ii) definition of badges and point awarding rules; and, iii) presentation of a leader board to inform on scores and to support awarding and management of points (by authorized evaluators). Total implementation costs amounted to only 12 programmer-hours. This low cost is evidence that the proposed gamified approach for learning ITST processes may be cost-efficient, thus providing a positive, partial-answer to the research question.

3 APPLICATION AND RESULTS

Application of the proposed game-based ITST learning approach was carried out at Connect (www.cnnt.com.br), a small software development company operating in mobile computing in Brazil since 2002. In this case study, Connect's ITST team was to carry out the migration of a sales support IT service from Palm OS to Android devices. Migration was for the benefit of a client's business units distributed over a wide operations area, from industrial to wholesale sectors (each "unit" comprised a number of users). The design and transition process for the new service involved evaluation of the Palm OS system legacy requirements; (re)implementation, testing, and validation of these requirements on the new Android platform; reconfiguration of the cloud hosting service to operate with both old and new systems during the transition phase; deployment; sales force training; training of other sales process stakeholders (supervisors, billers, IT staff and company managers) and finally, launch (go live). Due to particulars of each unit's requirements, all these tasks were evaluated prior to their execution to mitigate risks and to establish each unit's deadline for service launch (operation).

Connect's ITST team consisted of two software engineers, a trainee and a team manager, all male, aged 27, 24, 18 and 24, respectively. These actors joined the learning game voluntarily. The process of changing the actors' work routine started with training sessions on goals, objectives, and their linkage to desired outcomes; target ITSM BPs; BPM basics and notation; and, instructions on how to read and interpret the BPMN notation. A first version of the proposed "progression game", was presented to the team with the initial version of the BPD constructed for this specific transition case (mobile sales system). The point and badge award criteria were then spelled out besides introducing the persons responsible for evaluating the work and for granting rewards.

Application of the gamified ITST BPs evolved along three stages. In the first stage, Connect's ITST team operated using mostly undocumented, albeit ITIL-oriented deployment BPs which it felt, would likely lead to a high number of cases of perceived low quality deployment, rework to fix problems, and schedule overruns. In short, an overall, unsatisfactory client experience might ensue. The (pre)existing transition process was used for a month. During that time, the team became aware of a lack of established procedures to handle unsuccessful internal tests or validation by the client. In many cases, due to the lack of clearance about what needed to be done in the occurrence of these exceptional cases; changes were also carried with no documentation and without running all the needed

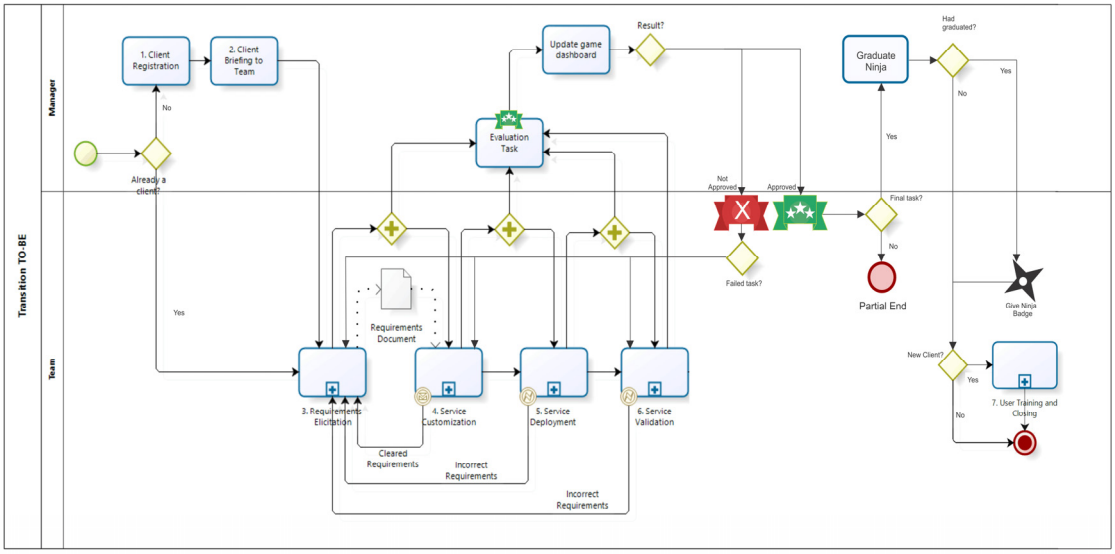


Figure 1:Illustration of gamified ITST BPD (Business Process Diagram).

tests over again. These findings offered the first opportunity to change existing ITST BPs and see how the change would be absorbed by the now gamified, transition team.

In stage two, which lasted 2 months, the process was refactored to include some exception-handling procedures to address the identified deficiencies. BPD changes were documented, explained to the team and incorporated in a new version of the transition BPD.

A few penalties were applied to the actors for not following the new procedures, but the new process as a whole was adopted smoothly. Two months into stage two, some difficulty on performing the deployment task was detected. This was generated by the deployment package building activity, a highly complex activity and that was not properly detailed in the process documentation.

In stage three, this activity was then broken up generating a new sub-process. The inclusion of this sub-process was the third major change to the gamified transition BPD map. ITST team's resistance to learning and absorbing this last change was higher than before. We attribute this resistance to the fact that the change, now, was not a team's suggestion but resulted from an issue detected by the manager's analysis of their work performance while deploying the new mobile sales system. Even with the loss of points, some weeks for training, coaching and monitoring were required until usage of the process stabilized. The final ITST BPD with sub-processes and game elements which supported the game-based learning ITST approach is illustrated in Figure 1.

The bands in Figure 1 represent the status of the performed task: the red "X" band is a task that failed approval by the manager and the green starry band represents an approved task for which points may be awarded. In the case of a final task in the process, the graduation evaluation will be executed and, if the actor reached the required amount of points to advance to the next graduation level, he would be awarded a ninja badge in the "shuriken blade" step of the process.

3.1 Evaluation

The evaluation of the gamified ITST BP learning approach was done both in quantitative and qualitative terms.

3.1.1 Qualitative Evaluation

The qualitative evaluation was carried out by structured and open interviews involving all team members and corporate stakeholders (shareholders). The structured interviews allowed for Likert Scale (Uebersax, 2006) answers only and were meant to capture the impressions and reactions of the transition process staff (except for the manager who had coauthored the proposal) to the following aspects:

- General ITST process vision: to determine whether the actors view the transition process (through the transition BPD) as a useful tool for orientation when performing ITST tasks and

whether the gamified process could effectively solve previously existing ITST problems;

- Professional confidence: to understand whether the gamified approach truly brought the ITST team members clarification, confidence about the tasks to be executed and also, whether the approach is of any help in uprooting any useless tasks in the process and whether the tasks led to relationship difficulties with the clients;
- Impressions on the gamified approach: basically, to probe whether the motivational triggers we designed the gamified approach to include were actually activated.

Results for the qualitative evaluation by the transition team´s members where obtained from the Likert Scale (1-'Totally Disagree"; 2-"Partially Disagree"; 3-"Neutral"; 4-"Partially Agree"; 5-"Totally Agree") as follows:

A. General process vision
   i. Assertion: "Upon receiving a transition task, I identify which tasks I need to perform by looking at the transition process diagram" – Answers: three "5".
   ii. Assertion: "I could deploy the mobile sales service on any kind of client following the transition process" – Answers: one "3", one "4" and one "5".
   iii. Assertion: "The number of problems I face in the transition process today, is significantly smaller than when the game-based ITST learning approach was not used" – Answers: two "4" and one "5".

B. Professional confidence
   i. Assertion: "Even with the gamified process at hand, I'm still confused about the tasks I should perform" – Answers: one "1" and two "2".
   ii. Assertion: "The transition process has dispensable tasks" – Answers: two "1" and one "3".
   iii. Assertion: "I faced problems with my clients due to the mandatory execution of a task in the transition process" – Answers: three "1".
   iv. Assertion: "Execution of the transition process tasks avoids problems during deployment" – Answers: three "5".

C. Impressions on the gamified approach
   i. Assertion: "I use the gamification points to evaluate my performance as an IT service transition analyst" – Answers: one "1" and two "5".
   ii. Assertion: "Being a transition 'Ninja' is important for me" – Answers: one "1" and two "5".
   iii. Assertion: "I seek to reach the "black belt" (Transition Ninja! Badge)" – Answers: two "3" and one "5".
   iv. Assertion: "I appreciate receiving bonus points notifications by e-mail" – Answers: one "1", one "4" and one "5".

It is important to notice that Assertion "A.i" relates to the research question in the sense that it guides the staff member to the next activity quickly (thus contributing to time-efficiency) and Assertion "A.ii" by helping transition staff to deal with complexity (thus contributing to effectiveness).

Evaluating the answers obtained in the interviews, we can see that a general process vision was indeed established in the minds of the actors. They clearly recognize the transition process as a valuable tool to resolve most of the problems they faced in the transition phase of this specific (sales automation) software lifecycle. The ITST actors also indicate that the game-based "learn by doing" method can help deliver a much more reliable service to their clients, therefore increasing their confidence on their own expertise.

As for impressions on the gamification aspects per se, answers diverged somewhat. The actor who is also a game design / development expert disagreed with the others on the motivational triggers we sought to achieve in three of four of the questions, remaining neutral in one of them. His answers possibly indicate that game-based learning approaches will not make much of an impression on teams with much technical expertise on game design. On the other hand, the other actors seemed to be positively impressed by all the triggers with exception of one specific case where the introjection seemed to have a neutral effect. We attribute this to a low intrinsic motivation for this one actor on becoming a more qualified transition analyst, possibly pointing to a lack of alignment between the work he had to perform and the professional objectives he set for himself.

In the open interviews the actors conveyed the impression that the game-based ITST learning approach helped them to speed up and materialize knowledge because it "connected and clarified everything that needed to be done" as commented by one and as another said: "the board game helped to remind what needed to be done next". (Although the approach still "needs more information on how to go about doing it".) Again, the impression of two players were along the line that "the gamified approach helped to create a competition atmosphere and pushed me to try to become "top ninja". The game expert however, felt this did not motivate

him, but he was "(pleasantly) surprised to receive point award e-mails".

Connect has three shareholders. One of them does not participate in the day-to-day operations of the company. The other two evaluated the gamified approach from the point of view of the company. The qualitative evaluation by the shareholders was carried out by an open-answer, five-question questionnaire which compared results after gamification as compared to before:

i. How do you evaluate the performance of the transition processes?
ii. How do you evaluate the assimilation of the transition process details?
iii. How do you evaluate the transition staff performance?
iv. How do you evaluate impact of the gamified approach in the transition cycle time and costs?
v. What was the impact of the gamified approach in client relationship?

The answers provided may be summarized as follows:

i. Improved transition activity flow. Improved inter-staff members and staff-client communication. Better Transition Process transparency, allowing all involved, including clients, to follow and anticipate steps.
ii. Faster learning curve. Higher productivity. Staff better acquainted with the ITST process.
iii. Improved performance. Less time spent on rework and trying to find out what to do next.
iv. Lower cycle time and hence, lower cost per transition.
v. Better informed client. Providing better input to transition staff. Lower rate of complaints.

The above answers provide clear, albeit preliminary, evidence in favor of a positive answer to the research question.

3.1.2 Quantitative Evaluation

The quantitative analysis was intended to focus on two simple but important measures to evaluate whether the gamified approach had some positive impact not only on Connect's employee motivation but also on the business itself:

- Time for full compliance with the new ITST process: by full compliance we mean that all process errors do not exceed 1% of the tasks performed in the whole process for every process execution in a period of three months;
- Average (whole) deployment process turnaround time: average of the time intervals from when the process is assigned to an actor to completion of (software) deployment. (Client-side delays were purged from evaluation figures).

Quantitative evaluation results indicate that full compliance with the new process was achieved four months after the last change to the ITST process.

Table 1 brings evolutionary milestones. Average turnaround deployment times for the sales automation system are given in Table 2 for some business units (BUs) before and after the gamified transition learning approach was adopted.

Table 1: Milestones.

Action	Date
First version	25th October 2012
Badges and Bonus pointing included	22nd November 2012
First badge achieved	29th January 2013
Final version	5th February 2013
Process Stability Reached	17th May 2013

Table 2: Deployment Turnaround Times.

Client		Complexity	Deployment Time
Before Game-based learning	BU 1	Medium	10 months
	BU 2	Medium	5 months
	BU 3	Low	6 months
	BU 4	High	3 months
After Game-based learning	BU 5	Low	1 month
	BU 6	Low	1 month
	BU 7	Medium	1 month
	BU 8	High	20 days
	BU 9	Low	18 days

"Complexity" in Table 2 indicates the level of deployment complexity as attributed by the team depending on the client's requirements and the number of business rules implemented or customized in the deployed sales system. Deployment time is the approximate time interval between a client's initial request and the final validation of the requirements by the client for a given business unit. One notices a significant reduction in deployment times in favor of the gamified approach. One should look into this reduction carefully, though, because it can be influenced by other factors such as diminished flow into other operational processes (like incident and

problem management which were performed in part by the same team); the team's growing technical expertise and experience with the ITST process and with the sales system technologies as time passed (which we believe made even more complex tasks faster); and, their disposition to learn.

On the other hand, Table 2 indicates an accelerated learning curve with the gamified approach. Assuming increasing expertise correlates with decreasing turnaround times, one can see that ITST expertise improves as time passes and with a rapid acceleration immediately after adopting the game-base learn by doing approach.

3.2 Validation

Gamification of ITST processes depends on the system of interest and it is on-going. Here, a game-based ITST learning approach was used by a small team of professionals. Therefore, the answer to the research question is preliminary and restricted to the above context of the case study. On the other hand, since this context is somewhat representative of the industry – particularly of small companies – the answer will be meaningful, at least in what concerns "face validity" (Litwin, 1995).

We say the proposed game-based ITST learning approach has face validity if it "looks like" it is going to lead to a positive answer for the research question.

To test the approach for face validity, we asked the participants in the case study to indicate what they thought the answer to the research question would be. The respondents, unanimously, gave "yes" as an answer (the corresponding Guttman scaling was "Yes", "No" and "Not sure").

4 CONCLUSIONS

This paper summarized research on a novel approach that blends gamification concepts, elements and tools to those of business process modeling (BPM) together with learning and teaching methodologies to communicate and deploy changes to IT service transition (ITST) processes. The approach was economically implemented and applied to the case of a small team charged with the migration of a mobile OS-based sales automation system in Brazil. Results indicate the team was able to learn and operate faster and more effectively with the approach.

By considering a small, IT service provider, the paper also provides early evidence gamification can bring benefits and be within technical and financial reach of firms in general, not just major IT players. That evidence, together with the focus on ITST, is the main contributions of this paper and endorses recent suggestions in the literature that ITIL gamification may offer a positive outlook for ITSM practitioners. Further work is needed to answer the posed research question with greater confidence, for different transition scenarios and for team compositions and sizes.

ACKNOWLEDGEMENTS

The authors thank anonymous CSEDU 2014 reviewers for their comments and suggestions.

REFERENCES

Arnold S., Fujima J., and Jantke K. P., Storyboarding Serious Games for Large-scale Training Applications. CSEDU (2013).

Bouvier P., Lavoué É., Sehaba K., and George S., Identifying Learner's Engagement in Learning Games: a Qualitative Approach based on Learner's Traces of Interaction. CSEDU (2013).

Castellani, et al. Game Mechanics in Support of Production Environments. Ext. Abstracts CHI'13, ACM Press (2013).

Deci, E. L., and Ryan, R. M. The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11, 227-268.

Deterding, S., Björk, S.L., Nacke, L., Dixon, D. and Lawley, L., Designing gamification: creating gameful and playful experiences. Ext. Abstracts CHI'13, ACM Press (2013).

Deterding, S., Sicart, M., Nacke,L., O'Hara, K., and Dixon, D., Gamification: Using Game Design Elements in Non-Gaming Contexts, Ext. Abstracts CHI'13, ACM Press (2013).

Kankaanranta, M. and Neittaanmki, P. Design and Use of Serious Games. Engineering (Vol. 37, p. 208). Springer (2008).

Litwin, M., How to Measure Survey Reliability and Validity, Sage Publications, (1995).

Macfarlane, I., An Introductory Overview of ITIL ® V3. itSMF (UK) Ltd, (2007).

Rosemann, M., Handbook on business process management 1: Introduction, methods, and information systems. Springer (2010).

Uebersax, J., (2006) Likert scales: dispelling the confusion. Statistical Retrieved from http://john-uebersax.com/stat/likert.htm.

Werbach, K. & Hunter, D., For the Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press (2012).

Bringing Real-life Practice in Software Project Management Training Through a Simulation-based Serious Game

Alejandro Calderón and Mercedes Ruiz
Department of Computer Science and Engineering, University of Cádiz, Cádiz, Spain
alejandro.calderon@uca.es, mercedes.ruiz@uca.es

Keywords: Software Project Management, Simulation, Serious Games, Bloom's Taxonomy.

Abstract: Nowadays, we can observe a lack of tools that allow teaching and assessing in Software Project Management in a more real-life way. In this paper, we discuss the need of developing tools to teach in the scope of Software Project Management, and how we can achieve that learners test their knowledge and cover the needed educational objectives, in a risk-free environment through a serious game. Hence, we propose ProDec, a simulation-based serious game to teach and assess in software project management. This tool is an attempt to surpass the limitations found in similar proposals. The paper also maps the stages of the game lifecycle to the levels of Bloom's taxonomy to show how the game helps to achieve different levels of educational objectives.

1 INTRODUCTION

After the software crisis in the beginning of the 1970s, experts began to be aware about the importance of software engineering as a means to provide a set of methods, tools and procedures for the development of quality software, within the constraints of cost and time. However, the 2011 edition of the CHAOS report found that only 37% of all Information Technology (IT) projects succeeded in that they were delivered on time, within budget, with all the required features and functions. IBM research on the reasons for IT project failure concluded that 54% of IT projects failures are a direct result of poor project management.

The importance of teaching software project management for IT learners has always been supported by organizations such as the Association for Computing Machinery (ACM) and IEEE-Computer Society in their joint task force curricula. In the Curriculum Guidelines for Undergraduate Degree Programs in Computer Science curricula recently released (ACM/IEEE, 2013), these organizations have not only highlighted the importance of this matter, but they have also emphasized the need of teaching software project management in a highly practical way, where learners can test their knowledge in real-life scenarios. By a highly practical learning, it is intended that future practitioners acquire professional practice *during*, and not after, their studies.

However, despite the importance of training in software project management, many authors conclude that software project management subjects are still basically taught following a highly theoretical pattern and, as a consequence, learners do not show much interest in them (Ibrahim, 2011).

Compared with other degrees, such as Medicine or Aeronautics, IT learners start their professional life with a serious lack of real-life practical skills. As a result, new professionals need to develop their experience by working in real projects, where the effects of an inadequate plan or a bad decision can lead to project failure or the loss of significant profit.

A serious game is a game with the purpose of training or educating users. They can help in situations like the one described above, as tools to acquire experience and motivate learners, given their engaging nature. Moreover, simulation-based serious games allow us to bring real-life scenarios into the learning process in a risk-free environment.

In this paper, we extend a previous work on ProDec (Calderón & Ruiz, 2013), which is a simulation-based serious game for software project management training.

In the scope of this work, we address the following research questions:

- RQ1: What are the main weaknesses of current serious games for teaching software project

management and how can ProDec overcome them?

- RQ2: What is the educational effectiveness of ProDec according to Bloom's taxonomy?

In order to answer these questions, we have performed the following steps:

1. We have searched and analyzed the related work on different proposals of serious games in software project management.
2. We have defined a set of criteria to compare these proposals and identified their strengths and weaknesses.
3. We have evaluated ProDec using this same set of criteria.
4. We have mapped the educational requirements defined for each level of Bloom's taxonomy with the different stages of ProDec lifecycle to find out the educational effectiveness of ProDec according to this well-known taxonomy.

This paper is structured as follows: Section 2 presents the works related to our proposal and provides a comparison of similar proposals Section 3 briefly reviews Bloom's taxonomy since that framework has been used to evaluate the coverage of learning objectives of our proposal; Section 4 describes the serious game developed; Sections 5, 6, and 7 show the evaluation of our proposal using Bloom's taxonomy; Section 8 provides information about the state of ProDec's evaluation process. Finally, Section 9 summarizes the paper and presents our conclusions and future work.

2 ANALYSIS OF CURRENT PROPOSALS

There are several serious games in the field of software engineering education as Caulfield, Xia, Veal, and Maj show in their systematic review of the literature (Caulfield et al., 2011). However, this study also shows that the serious games that focus on software project management are scarce and have a quite specific scope. If we focus on the serious games developed for the area of software project management, the following ones are well-known examples:

- SIMSOFT (Caulfield et al., 2011) is a serious game materialized as a printed game board, that shows the players the flow of the game, and a Java-based board, where the players can see the current and historical state of the project and adjust the project's settings. It mainly focuses on human resource management, with an emphasis on how the ability of the staff affects the outcomes of the project.
- DELIVER! (Von Wangenheim et al., 2012) is also based on a printed game board designed to help learners develop the skills needed to measure and control project performance by applying the Earned Value Management technique (Project Management Institute, 2005). One of the main aims of this game is to motivate students in their learning process.
- SimSE (Navarro & Van Der Hoek, 2004) is a serious game completely developed as a software tool that is based on software project simulation. The game supports several development methodologies and focuses on the development of abilities for software process management.
- SESAM (Drappa & Ludewig, 2000) is another serious game developed as a software application that uses simulation techniques to motivate learners in learning software project management. Players take the role of a project manager and must plan and control a simulated project.
- The Project Manager Game (Games by Robc, 2013) is an online serious game where users have to allocate the most suitable staff to particular tasks and complete their project on time within the allocated budget.

In order to perform a comparison of these proposals and also to identify their strengths and weaknesses, we propose the following set of features. These features are based on our analysis and also on the ones used by Caulfield and his colleagues (Caulfield et al., 2011).

a) F1. Coverage of software project lifecycle, that is, the phases of the project lifecycle the game deals with.
b) F2. Coverage of the Project Management Body of Knowledge (PMBOK) (Project Management Institute, 2013), that is, the different processes and techniques described in the PMBOK that the game helps to learn and practice with.
c) F3. Support of revised Bloom's taxonomy (Krathwohl, 2002), that is, the level of coverage of each level of Bloom's taxonomy the game offers.
d) F4. Support for automatic learning assessment, that is, the level to which the game helps trainers to assess the learners automatically.
e) F5. Simulation-based, that is, if the game makes use of simulation techniques, and to which extend.

f) F6. Game flexibility, that is, if the scenarios of the game are static or they can change dynamically during the game play.

Table 1 summarizes the comparison of the serious games previously described. In this table, the columns represent the serious games analyzed and the rows are the set of features already defined

According to the PMBOK, created by the Project Management Institute (PMI), the project lifecycle is defined by five stages: Initiation, Planning, Controlling & Monitoring, Executing and Closing. Those stages give also name to the five groups of management processes that need to be carried out with the help of techniques. All the serious games analyzed focus on a specific stage of the project lifecycle (F1) and/or help to practice a specific process or technique in that stage (F2). None of the games analyzed give learners the possibility of studying the complete lifecycle of a project from the initiating stage, where the project characteristics are defined, to the closing stage, where the project ends and it is the time for analyzing the results and deriving the lessons learnt.

If we compare the educational objectives that can be achieved by using these games with a well-known taxonomy of learning objectives such as Bloom's taxonomy, we can conclude that only SIMSOFT covers all the levels of the taxonomy.

Regarding the assessment of the skills that learners acquire by using the games (F4), all the games analyzed use traditional methods for learners' assessment, and none of them includes an automatic process of gathering and analyzing information about the game play to support an automatic assessment. Furthermore, only two out of the five games analyzed make use of simulation as a means to add realism to the game scenarios (F5). However, the two games that make use of simulation provide scenarios supported by static simulation models. This means that the scenarios the player can play are fixed offering a poor flexibility and leading to an early loss of player's motivation.

From the analysis of the current proposals of serious games for teaching software project management, we found that their main weaknesses are:

- They focus on learning specific techniques of project management or specific stages of the project's lifecycle.
- They do not usually reach all levels of Bloom's taxonomy.
- They do not allow to assess learner's new skills automatically.
- They do not offer flexibility.

Taking this analysis into account, we can conclude that further research is needed to overcome the weaknesses found in the application of serious games for software project management. For this reason, in this paper, we propose ProDec, a serious game designed to overcome these weaknesses.

Table 1: Comparison of serious games for software project management.

Serious Game	SIMSOFT	SESAM	SimSE	DELIVER!	The Project Manager Game
Coverage of software project lifecycle (F1)	Planning, Controlling & Monitoring	Planning, Execution, Controlling & Monitoring	Planning, Execution, Controlling & Monitoring	Monitoring	Planning
Coverage of PMBOK (F2)	Staff management	Staff management	Staff management	Earned Value Analysis	Staff management
Coverage of Bloom's taxonomy (F3)	All levels	Knowledge	Knowledge	Application	Application
Support for automatic assessment (F4)	Not allow to obtain automatic assessment reports				
Simulation-based (F5)	No	Yes	Yes	No	No
Game flexibility (F6)	Statically scenarios provided by the game			Static board scenario	Random scenarios

3 BLOOM'S TAXONOMY

Bloom's Taxonomy (Bloom et al, 1956) is a widely accepted classification of learning objectives within education. It constitutes a common framework for learning proposals' assessment and comparison. In the world of serious games, this taxonomy is also used for these aims. For this reason, we have also chosen Bloom's taxonomy to assess the educational objectives that our proposal can reach.

Bloom's taxonomy refers to a classification of the different objectives that educators set for learners. It divides educational objectives into three domains: Cognitive, Affective, and Psychomotor. If we focus on the Cognitive domain, where skills revolve around knowledge, comprehension, and critical thinking on a particular topic, we find six levels sorted in a hierarchy. These levels, in ascending order, are shown below:

- Knowledge, the subject is able to recall previously learnt information. They recognize information, ideas, facts, dates, etc. in an approximate way as they have learnt.
- Comprehension, the subject is able to demonstrate the understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating the main ideas.
- Application, the subject is able to solve problems in new situations by applying acquired knowledge, facts, techniques and rules in a different way.
- Analysis, the subject is able to examine and break information into parts by identifying motives or causes, and they are able to make inferences and find evidence to support generalizations.
- Synthesis, the subject is able to create, integrate, combine ideas, pose and propose new ways of doing. They are able to apply previous knowledge and skills to produce something new or original.
- Evaluation, the subject is able to make judgments about information, validity of ideas or quality of work according to personal opinions based on a set of criteria.

A revision of Bloom's taxonomy (Anderson & Krathwohl, 2001) establishes that in the cognitive domain there are six levels but not all are sequential in a hierarchical order. This revision proposes that the three lower levels are sorted in a hierarchical way, but the other three, the higher levels, are placed in parallel levels.

Like Caulfield, Xia, Veal, and Maj show in their serious games review (Caulfield et al, 2011) that the majority of authors in this scope are satisfied with placing their educational objectives at the basic levels of Bloom's taxonomy, even some of the proposals only reach the Knowledge one. The only exception that we found is SIMSOFT that reaches the higher levels of the taxonomy.

In the following sections we describe the different stages of ProDec and we perform an analysis of Bloom's taxonomy levels that our proposal satisfies.

4 PRODEC OVERVIEW

ProDec is a serious game to teach and assess learners in software project management. Its main goal is that learners acquire decision-making skills on problems that can appear within a software project lifecycle. Consequently, it helps learners start their professional career with some practical abilities for their profession.

The aim of the game is to successfully manage a software project. The game is over when the project significantly overruns either the approved budget or the allocated time. The player wins when they are able to complete the project within the time and cost limits. Besides, ProDec not only focusses on problem solving during the executing and controlling stages of a software project, but it also invites players to fully plan their own project and then monitor and control its progress by simulating its execution.

At the same time, ProDec helps trainers in the assessment of the skills that learners must acquire by playing the game. To do this, ProDec accepts the assessment criteria as an input of the trainer, and automatically: a) gathers data about these criteria during the game plays, b) analyzes the data collected, and c) shows an assessment report for the learner and the trainer.

ProDec is a game to be played in teams, so that it can also help to develop some soft skills in project management such as leadership and communication skills. Actually, ProDec does not teach learners the basic principles of software project management. Before playing, the players need to acquire these principles in lectures. Therefore, ProDec is a tool to be used in advanced stages of an academic course. Although the game is thought to be used by teams, it can also be played individually.

4.1 Lifecycle

ProDec offers players two ways of playing, a full game and a quick game. In a full game, players begin a project from scratch while, in a quick game, players can select one of the scenarios previously uploaded by the trainer. In this second case, players can only practice their project monitoring and controlling skills. Disregarding the mode of playing, players go through different stages of the project lifecycle with the goal of ending a project successfully. Within the game, there are three phases of a game play's lifecycle, which are Onset, Execution and End phases.

The following sections further describe the functionality associated with each phase of the game and the level of coverage of the learning objectives according with Bloom's taxonomy.

5 ONSET PHASE

ProDec's Onset phase is the first contact that learners have with the game. If the players select to play a quick game, they have to go through the different project scenarios available in the game and read and evaluate each project's features. In this case, learners need to know the main concepts and principles about software project management and need to understand the information and data that the game shows, so that they can get an idea about the difficulty of each proposed project scenarios. So, in this phase, learners need to *remember* and *comprehend* the information provided about the software projects to begin the play successfully. In this way, Knowledge and Comprehension levels of Bloom's taxonomy are covered.

On the other hand, if players select to play a full game, they follow a process that guides them in making the software project plan. This process is made of five sequential stages and allows learners to provide all the data needed to create a new software project plan. The stages that make this process are the following: project information, size estimation, project team definition, tasks definition and risks analysis.

5.1 Project Information

Project Information is the first stage of the process of creation of a software project plan. In this stage, learners have to enter the general information of the project about its scope and features, such as the salary of the workers, the length of the project, the numbers of use cases, etc., that are necessary to begin the size estimation stage. In order to do this task, learners need to know and understand the different concepts about software project information that the game uses. If learners do not know or understand properly these issues, it is likely that they will end up in a failed game, as it happens in real life, where a project misinterpretation leads to a failed project. The training about the concepts used in the phase of the game has to be offered during the lessons taken before playing. As in this stage, in all the next stages of this process learners need to know, properly understand and remember the concepts learnt in the lectures, so that the data they provide to the game is consistent. Hence we can see that Knowledge and Comprehension levels of Bloom's taxonomy are covered by the game.

5.2 Size Estimation

In this stage, players provide the estimate of the size of the project starting with the total number of use cases. Then, learners need to calculate and enter the size of each use case. To estimate the size of each use case, learners use Albrecht's Function Point Analysis (Albrecht, 1979). A function point is a unit of measurement to express the amount of functionality of an information system. In this stage of the game, learners need to *apply their knowledge* to calculate the different size of the use cases before entering the data to continue. Consequently, the Application level of Bloom's taxonomy is covered.

Besides, in this stage learners apply Albrecht's technique and get the results of under- or over-estimating in real-life project scenario, which is a different situation to the traditional one focused on blackboard activities.

5.3 Project Team Definition

During this stage, learners design their project team by defining its members. For each team member, players have to select some features for their personality and past work experience. Currently, the personality of an employee is made of two traits that the player has to select from a range of twenty two available ones. The set of personality traits included in the game is based on the sixteen personality factors described by Cattell (Cattell et al. 1988). In this stage, the game helps learners to think about the different outcomes of mixing personalities in a team and the importance of achieving a good team synergy in a successful project.

5.4 Tasks Definition

Once the size has been estimated and the project team designed, the following stage asks players to define the project tasks. In this stage of the process, players define the project tasks, and enter, for each of them, the time data, the budget allocated, and its predecessor tasks. Consequently, this stage asks the player to provide the information gathered in a PERT diagram (Moder, 1983) of the project. PERT technique is recommended in PMBOK's Time and Cost knowledge areas and it is included in the Project Management Professional (PMP) exam. This function allows learners to apply their knowledge about PERT diagrams in real-life scenarios and to analyze its features within a project unlike more traditional approaches where knowledge is exclusively acquired by solving individual exercises during the lectures. Besides, in this stage, players have also to allocate tasks to the team members. To do this successfully, they need to analyze all the information entered about the tasks description and the personality traits and work experience of the team members with the aim of allocating the most suitable staff to each task. It can be seen that the Analysis level of Bloom's taxonomy appears in this stage of the game.

5.5 Risk Analysis

The last stage of the process that players follow to make the software project plan is the risk assessment and analysis. As its name indicates, in this step, the players perform the quantitative analysis of the risks that can appear along the execution of the project. In a quantitative risk analysis, players need to enter the risk probability and the loss magnitude for every risk identified. The risk probability is the likelihood of occurrence of a risk. The loss magnitude is the potential loss the project may suffer in the case that the risk appears and it is not properly controlled. Later, these risks will be transformed into events of the simulation model. The probability that these risks occur is defined by the risk probability. The loss magnitude establishes the time delay that the tasks being performed at the moment when the event occurs suffer.

Based on the knowledge acquired in the lectures and the information entered about the project, the players are required to analyze the project within its scope with the aim of making a proper list of the risks to be taken into account. Again, we can see how the Analysis level of Bloom's taxonomy is covered.

Along all the onset phase, learners have to analyze all the elements of each stage to get information and synthesize this information into new ideas and decisions to make. Moreover, given the social character of the game, before making a decision such as allocating staff to a task or defining the probability of a risk, players need to analyze the information of the project to argue, defend, discuss, evaluate, negotiate and agree the best decision within their team to make a good project plan. This feature allows the learner to work at the highest levels of Bloom's taxonomy which are Synthesis and Evaluation. Hence by playing a full game, ProDec provides a full coverage of Bloom's taxonomy.

6 EXECUTION PHASE

The second step consists on executing the project created or selected in the previous phase. To support this project lifecycle phase, ProDec automatically generates a source code file with the equations of a discrete-event simulation model that simulates the project described in the first phase of the game. In addition, the file also contains the source code of the user interface for the specific simulation model generated. It is important to notice that this feature helps ProDec achieve a high level of flexibility given that the number of different projects that can be simulated is unlimited.

To achieve this flexibility, we have performed a reverse engineering process. We have analyzed how the simulation software used implements the elements of a discrete-event simulation model, and then, we have generalized the procedure with the aim of building the simulation models in an automatic way. Working like this, ProDec is able of simulate any project plan the players create overcoming the lack of flexibility found in other proposals previously described.

Once the source code of the simulation model is generated, the simulation model is launched and the players start managing the project. During the simulation of the project execution, the game shows the players several screens where the progress of the project is presented and different actions are provided to control the project. In this phase, learners practice two main concepts. On the one hand, they put into practice their knowledge about the Earned Value Analysis (EVA) for monitoring the progress of the project.

On the other hand, the learners practice their decision-making skills by correcting the potential

deviations of the progress of the project from the goal of ending the project within the time and budget. If a corrective action is needed, players select it from a set of actions such as hire a team member or reorganize the project team.

In this phase of the game, players need to analyze the information presented in different screens for controlling and monitoring the progress of the execution of the project. For this, learners use their knowledge about software project management to understand the information about the progress of the project provided by the game. They also have to analyze, create, discuss, argue, evaluate and negotiate within their team to agree on the best decisions in each moment. In this stage, the game also presents a full coverage of the six levels of Bloom's taxonomy.

7 END PHASE

The last phase consists on the players' assessment. By using the information that ProDec has been recording during the game play and the assessment criteria established by the instructor, ProDec generates an assessment report of the learners describing their level of achievement. These information records come from several sources within the game such as the project plan with the initial estimates, the project monitoring data and the kind of decisions that the players made during the play.

The assessment criteria are provided by the instructor in the form of a rubric. An assessment criterion links the information recorded in the rubric with the information recorded during the game. By using a labelling system the labels describing the skills of an assessment criterion are matched with the records of the game that contain the information needed to assess such criterion.

As a consequence, ProDec is able to perform the learners' assessment by analyzing the information recorded during the game and applying the assessment criteria set by the instructor, concluding with the generation of a detailed assessment report that describes the skills acquired by the players.

The generated report allows players to analyze the course of the game play with the aim of learning about their experience. At this point, learners can learn from their own mistakes and can analyze the events occurred along the game to get new knowledge and generate new ideas for future plays.

In this stage, the highest levels of Bloom's taxonomy are also covered.

8 CONCLUSIONS

In Section 2, we presented a comparison between several serious games attending to a set of features considered to be of importance in a game for software project management training. We showed that none of the analyzed games fully satisfies all the mentioned features. In this paper, we proposed ProDec, a simulation-based serious game for software project management, which aims to teach, assess and motivate learners in learning and practicing the principles of software project management as well as improving some important soft skills, like project leadership, in a risk-free environment provided by project. Thus, learners can experiment their abilities with real life scenarios without costs or risks.

ProDec has been developed with the aim of having a tool that satisfies all the criteria that other serious games in the field do not cover. Therefore, ProDec is a serious game that:

a) Opposite to the majority of the serious games in the field, which are focused on a specific stage of the project lifecycle or the training of a specific technique, allows learning along the whole project lifecycle. To achieve this, players can start a software project plan from scratch, improve their skills in decision-making, analyze the log of games played in the past, etc. ProDec offers a sufficient coverage of all the levels of Bloom's taxonomy, from the Onset stage until the End stage of the game lifecycle.
b) ProDec also allows trainers to assess the learners' skills through the information reports that the tool generate during the game play. Besides, at the end of a play, ProDec automatically provides an assessment report with detailed information about the events occurred in the game. Learners can then analyze their game experience to achieve new knowledge and improve their abilities in software project management.

Finally, ProDec is a simulation-based serious game. This means that it uses simulation to execute the software projects, planned by learners and instructors, so that players can practice decision-making in the scope of project controlling and monitoring processes. As a remarkable feature, ProDec generates in real time the source code of the specific discrete-event simulation model for the created or selected project together with the appropriate user interface. This feature makes ProDec a completely flexible tool regarding the

unlimited project scenarios that can be simulated, and, therefore, played.

To sum up, along this paper we have shown how ProDec is a flexible tool, covers the levels of Bloom's taxonomy, allows players to take contact with all the lifecycle of a project and helps professors to assess learner's skills, overcoming the weaknesses identified in current proposals, answering to our first research question. Besides, we have analyzed how ProDec reaches the six levels of Bloom's taxonomy by analyzing the player's behavior ProDec demands at its different lifecycle stages, answering, this way, our second research question.

Our aim is to create a tool to support the effective practical training of the processes of software project management. We believe that this kind of tool is needed to prepare the new practitioners for their professional life in the best possible way. For this reason, we are currently working to improve the features of ProDec to increase learners' soft skills such as project leadership, motivation, engagement and competitiveness. To achieve these goals, we are also studying the benefits of integrating our game with social networks and adopting gamification strategies for enriching the learning process.

Moreover during the first semester of 2014 the tool is being evaluated with real learners. The lessons learnt from this process will be used to measure the effectiveness of the learning and assessment processes when using ProDec and to improve the game.

ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Ministry of Science and Technology with ERDF funds under grant TIN2010-20057-C03-03.

REFERENCES

ACM/IEEE-CS, 2013. Computer Science Curricula 2013.
Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. December 20, 2013 The Joint Task Force on Computing Curricula Association for Computing Machinery (ACM) IEEE Computer.

Albrecht, A. J., 1979. *Measuring application development productivity*. Proceedings of the joint SHARE, GUIDE, and IBM Application Development Symposium, Monterey, California, IBM Corporation, pp. 83-92.

Anderson, L. W., and Krathwohl, D. R., 2001. *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives.* Longman, New York.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R., 1956. *Taxonomy of educational objectives: the classification of educational goals*. Handbook I: Cognitive Domain. New York, Longmans, Green.

Calderón, A., and Ruiz, M., 2013. *ProDec: a serious game for software project management training.* The 8th International Conference on Software Engineering Advances. ICSEA, 565-570 Venice. Italy.

Caulfield, C., Veal, D., and Maj, S.P., 2011. *Teaching software engineering project management-A novel approach for software engineering programs*. Modern Applied Science, 5(5), 87-104.

Caulfield, C., Xia, J. C., Veal, D., and Maj, S. P., 2011. *A systematic survey of games used for software engineering education.* Modern Applied Science, 5(6), 28-43.

Cattell, R. B., Eber, H. W., and Tatsuoka, M. M., 1988. *Handbook for the sixteen personality factor questionnaire (16PF)*. Institute for Personality and Ability Testing.

Drappa, A., and Ludewig, J., 2000. *Simulation in software engineering training.* Proceedings of the 22nd International Conference on Software Engineering, pp. 199-208. ACM.

Games by Rocs, 2013. *The project management game.* http://thatpmgame.com/ [retrieved: December, 2013].

Ibrahim, I., 2011. *Teaching project management to IT students: methods and approach*. 2nd International Conference on Education and Management Technology. IPEDR. Vol. 13, 185-191. IACSIT Press.

Krathwohl, D. R., 2002. *A revision of Bloom's taxonomy: An overview.* Theory into practice. 41(4), 212-218.

Moder, Joseph J., 1983. *Project management with CPM, PERT, and precedence diagramming*. 3rd ed. New York: Van Nostrand Reinhold.

Navarro, E. O., and Van Der Hoek, A., 2004. *SimSE: An interactive simulation game for software engineering education*. Proceedings of the Seventh IASTED, International Conference on Computers and Advanced Technology in Education.

Project Management Institute, 2013. *A guide to the Project Management Body of Knowledge (PMBOK® Guide)*. Fifth Edition.

Project Management Institute, 2005. *Practice standard for earned value management*.

Von Wangenheim, C. G., Savi, R., and Borgatto, A. F., 2012. *DELIVER! – An educational game for teaching earned value management in computing courses.* Information and Software Technology, 54(3), 286-298.

Blended Learning as a Strategic Method Against the Illegal Use of Internet

Julija Lapuh Bele[1,2], Andreja Sladoje Jemec[2], David Rozman[2] and Maja Dimc[2]
[1]*Faculty of Management and Law, Ljubljana, Slovenia*
[2]*B2 d.o.o., Trzaska cesta 42, Ljubljana, Slovenia*
{julija.bele, maja.dimc, david.rozman}@b2.eu

Keywords: Cybercrime Prevention, Information Security, Blended Learning.

Abstract: The article addresses the issue of strategic prevention and fight against cybercrime related to children and teenagers with the use of blended learning in order to establish greater awareness and increase the knowledge of children, teenagers as, parents and educators regarding illegal internet content and related activities. Based on the theoretical background, practical experience and the analysis of questionnaires, we have prepared blended learning courses for each target group, which aim to raise the awareness of stakeholders (i.e. children, teenagers, teachers, parents) and contribute to cybercrime prevention and victim assistance.

1 INTRODUCTION

The increasing integration of web technologies in everyday life together with the popularity of social networks and development of mobile technology contributes to the creation of an optimal environment for various types of cybercrime and illegal internet content. Children and teenagers represent the most avid users of new technologies and functionalities, while they are at the same time the most naïve. Moreover, the general public is not aware of the severity of the problem.

As a part of the General Programme on Security and Safeguarding Liberties, the European Commission established the Prevention of and Fight against Crime Programme in order to contribute to the strengthening of the area of freedom, security and justice. The project "Education as a Strategic Method Against the Illegal Use of Internet" is funded within the framework of this programme with the key goal of increasing the knowledge and awareness of the general public, children and teenagers in particular, regarding the issue of cybercrime and illegal use of Internet.

The project includes the development and dissemination of best practices in the field of protection and support of cybercrime victims. Existing initiatives in this field are primarily focused on presenting information in different forms; however, we believe that developing an educational module and actively implementing it in primary schools will contribute to increased awareness of both children and adults consequently resulting in:

- increased level of uncovered illegal content,
- faster and easier work of law enforcement agencies, due to greater knowledge of cyber victims and increased level of reporting the incidents,
- decreased level of the number of cybercrime cases due to the increased knowledge regarding information system security.

We believe that only active involvement of the target groups will produce effective results.

In the EU Kids Online survey, only one third of 9-16 year olds (33 percent) stated that their parents know more about the internet than they do (Ólafsson and Livingstone, 2013). We thus believe that all target groups need to substantially improve their knowledge regarding internet safety and the issues of illegal internet use.

The main aspect of our project is to develop and implement an innovative educational module that will result in active involvement of the target groups in the activities of prevention and fight against illegal internet content and related activities. Namely, the educational modules teach the target groups how to protect themselves, as well as how to appropriately react when faced with illegal internet use, with emphasis on raising the level of reporting to relevant agencies. Additionally, the project

addresses the issue of the protection and support of victims of these types of cybercrime not only through the establishment of clear guidelines and recommendations, but also with the establishment of an online cybercrime victim "hot-line".

2 CYBERCRIME PREVENTION

In addition to its various benefits, the internet also contributes to increased exposure to different forms of crime. Today, the Internet enables certain offenses, which were unimaginable in the past. The level of privacy has substantially decreased as the general public willingly publishes personal information, which, of course, once published can never be erased – the fact that the general public, children and teenagers in particular, seem to easily forget. Furthermore, fraud and scams are successfully abusing the virtual environment in which boundaries and time are irrelevant.

Young internet users should be aware of the threats to their identity and wealth, as well as potential future impact of their actions. Furthermore, the core rules of etiquette in the virtual environment should be communicated to all users. Through raising the level of awareness of our children, we will be moving toward the creation of an information security culture in the long run.

2.1 Information Security Issues

Information security is a very wide area. It encompasses both technical security, as well as the threats posed by the users themselves, whether that is due to the lack of knowledge or naivety when exposed to social engineering. In relation to the assurance of technical security, IT professionals can install firewalls, antivirus software and enable regular updates of the operation system and antivirus software. However, there is no software to protect the system from its weakest link – the human. As the success of a social engineer is the result of greediness, trustfulness, naivety and, especially in younger people, inexperience, protection against social engineering can be implemented only with the users' personal knowledge, attention and care.

Parents can select appropriate software to improve the security on the child's computer or mobile phone. However, unfortunately many parents do not have sufficient knowledge in order to ensure child's safe participation in the virtual environment.

When it comes to information security, we primarily consider PCs and often forget the mobile devices, especially smartphones and tablets. According to Gartner Inc. (Gartner Inc., 2013) the proliferation of lower-priced tablets and their growing capability is accelerating the shift from PCs to tablets. The number of smart phones is also growing rapidly. In January 2013 Center for Safer Internet Safe.si (Safe.si, 2013) published the results of the research among internet users, which suggests that parents consider the appropriate age when a child gets a mobile phone to be 11 years. Almost every teenager has a smartphone. The number of applications for mobile devices cannot be counted. Each user has installed many of them. However, in these devices, we do not pay enough attention to the issue of information security.

Regardless of the platform, the main sources of threats are the interactive functionalities of the modern World Wide Web, especially in relation to social networks. Creating a profile is a prerequisite for joining the social networks. The profile includes information about the user along with pictorial material. When a user creates a profile, he/she can search for other users, different links, and can collect and share contacts list. The use of social network systems is increasingly popular. Interest groups attempt to realize their interests by using such a medium. The common tendency that an individual collects a multitude of "friends" is not derived from the human need for companionship, but from the need for status. For many users it could be said that the whole point of these networks is in the accumulation of as many friends as possible (Rosen, 2007).

In addition to the positive characteristics and trends, such as integration of users with common interests and keeping in touch with real friends worldwide, the users can be quickly exposed to abuse, inconvenience, and invasion of privacy. After creating his/her own profile, the user often forgets that the web is like a "bulletin board" and that the information published on the internet stays there forever. It is practically impossible to trust all online friends as much as friends in the real world, though via online social networks the users often behave in the same way. The consequence of this is the fact that the increased use of online social networks excessively reduces the level of self-protection, while at the same time also decreasing the level of attention, thereby unduly increasing the level of trust (Gregoric, 2010).

Users of the internet should be aware of the risks and pitfalls, as they can ensure their own safety with caution and taking appropriate action. The set of traps and risks may include: poor passwords, spam,

fake websites, false internet phone calls, identity theft.

2.2 Cybercrime and Its Impacts on Young People

Children and adolescents represent a vulnerable group of users who spend a lot of time on the web and social networks. They are exposed to the same threats as seniors, some of which can be devastating. Due to its serious consequences, we have focused on the following forms of crime as it relates to young people:

- Cyberbullying - bullying of children and teenagers (threats, harassment, humiliation, embarrassment, etc.) carried out by children and teenagers with the use of internet, digital technologies or mobile phones.
- Online sexual harassment and grooming – includes all actions with the goal of lowering the child's inhibitions in order to sexually assault the child.
- Child pornography and the dissemination of inappropriate content - all materials showing children or teenagers in inappropriate sexual context.

Child development can be harmfully affected by improper internet content. Inappropriate content constitutes contents that foster violence, spread hate speech and racial intolerance, enable online harassment, humiliation and insults, mockery of bodily defects, distribution of pornographic material to children, encourage dangerous activities, show examples of extreme violence and last but not least allow playing video games with violent and other problematic content. Furthermore, websites such as Pro-Ana promote anorexia, publish photos of boney models and celebrities as an incentive for extreme weight loss called "thinspiration". They promote anorexia and bulimia as a lifestyle rather than as a disease or food disorder. This is a dangerous message to young boys and girls who are in a vulnerable period of adolescence and receptive to environmental influences.

Online pornography has become a global problem. UNICEF estimates that more than 4 million websites, which show juvenile victims and even children younger than two years, can be found on the internet (Cehovin, 2010). Such an experience causes long-term effects on the child's later life both in terms of psychological effects, such as feelings of guilt, responsibility for the abuse, low self-esteem, feelings of inferiority and depression, as well as in the light of the fact that a child is victimized every time anybody watches material depicting his/her sexual abuse. Moreover, due to the simplicity of spreading the materials with the use of web technology, it is very difficult to stop the abuse. It is impossible to completely remove the materials, once published on the Web (Dimc & Dobovsek, 2012).

2.3 Digital Identity

Digital identity is made up of the following four categories of information (Maurel, 2009):

- Authentication elements: email address, user name, password, last name, first name, alias, IP address, etc.
- Data: personal, administrative, occupational, banking, social data, etc.
- Identifiers: photograph, logo, image, avatar, etc.
- Digital traces: contributions to public content management systems such as Twitter, YouTube, Facebook, Wikipedia, etc.

Digital tracks are practically indelible. Internet users, especially children, are unaware that their digital identity may affect their lives. Furthermore, inappropriate comments and posts, inappropriate pictures and videos can lead to abuse and can negatively affect their everyday life.

Active participation on the Internet should be seen as something positive; of course, if a person creates a positive online identity.

2.4 How to Prevent Cybercrime among Youth

Nowadays, parents cannot and should not prohibit the use of online environment and activities on social networks. However, they can increase the awareness of the risks and teach the children how to protect themselves.

Children and teenagers, due to their avid use coupled with naivety, are in much greater danger of publishing too much personal information and creating too open-public profile in social software applications.

Due to the widespread use of social networking sites by children and adolescents, it is important for parents to play an important educative role. Their task is not prohibiting the use, but rather raising the awareness about the consequences of actions performed in the virtual environment. Complete control by parents is practically impossible, since children and adolescents can monitor and change their profiles at school, with friends, via mobile devices. Therefore, the main task of the parents is to communicate the awareness of the potential hazards

and teach their children how to appropriately react when faced with troubles online. Adults should supervise children and teach adolescents to recognize the bad in the virtual environment.

Parents and teachers should talk to the children in order to warn them of the risks posed by irresponsible behavior on the internet, and children should have a sense that they can always turn to them. Therefore, the parents should constantly improve their knowledge on the subject.

Parents and children ought to spend some time online together. Children should inform their parents about their favorite websites. Parents should also have control over the time a child spends on the Internet. Child's computer should be in a room accessible by all family members. It must be protected by a filter, which blocks inappropriate content. Parents should also emphasize the value of personal data. They must tell their children that personal data, photos, name, address, school or data regarding their family members should not be given without their permission. Parents ought to alert the children to be careful when faced with any proposal for a meeting from people they met online. Without the knowledge and permission of the parent or responsible adult, a child should know that he/she should not meet anyone, despite the promising potential of the meeting. Children need to be aware that the people they are talking to on the internet may not be who they present themselves to be (Spletno oko, 2013).

3 PRELIMINARY RESEARCH

In order to appropriately adjust the learning content included in the project to the current issues and existing knowledge of target groups, we performed a preliminary research dealing with the level of awareness regarding cybercrime and information safety. In addition to the performed analysis on a group of children, teenagers and adults, we also reviewed the existing research in the field of cybercrime and information safety in Slovenia, as well as EU in general. The findings of existing research, combined with the findings of our preliminary analysis, were consequently used for the selection of the key areas to be included in our educational modules.

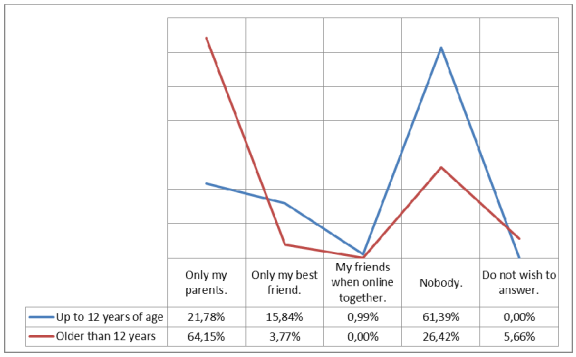
The analysis was performed with the use of two questionnaires intended for children and teenagers. The first questionnaire was prepared for children up to 12 years of age the second questionnaire was prepared for teenagers (older than 12 years). The first questionnaire contained fewer questions than the second questionnaire, since we assumed that the children in the first group are only beginning to use the internet, while the teenagers are already using a wider variety of internet functionalities. The third questionnaire was prepared for adults (parents and teachers), since we believe it is crucial for parents and teachers to be included in the educational process regarding cybercrime and information safety. Parents and teachers perform a crucial role in the child's entrance to the virtual world; namely, they should educate the child regarding the positive and negative effects and not merely rely on restriction and control. For the distribution of the questionnaire, the chain-referral sampling method was used; the questionnaire was given to a selected group of participants, which then recruited future subjects among their acquaintances. The total number of participants included 205 adults, 132 children under the age of 12 and 128 children older than 12 years of age.

The children and teenagers included in the analysis all use the internet and only 0.1 percent of the participants stated that they do not have home internet access; however, they use the internet in school. Their answers were compared to the rest of the participants and we found no differences in the level of use and awareness.

The comparison of the answers between children and teenagers displayed differences primarily in their relation toward parents, namely, older children are less likely to turn to their parents when in doubt of when in trouble, which is a reason for concern, especially due to the fact that teenagers spend large amount of time on the internet. The latter can be attributed also to the increasingly wider use of smartphones.

One of the first steps toward securing your information online is undoubtedly the use of passwords. Therefore, we were interested if the children and teenagers are aware of the importance of safekeeping their passwords. The results were encouraging, since 60 percent of the children stated they do not share their passwords with anyone. However, concerning is the fact that the percentage of children who would entrust their friends with their password increases with the age of the child, namely the percentage is very low in case of the children under the age of 12 (approx. 4 percent) and raises sharply with older children (approx. 16 percent). Similar were the findings related to the question of safeguarding privacy on the internet, personal data in particular. We found that younger children are much more likely to consult with their

parents (90 percent) before posting personal information, while this percentage drops with the increasing age of the child (58 percent in the group of teenagers).

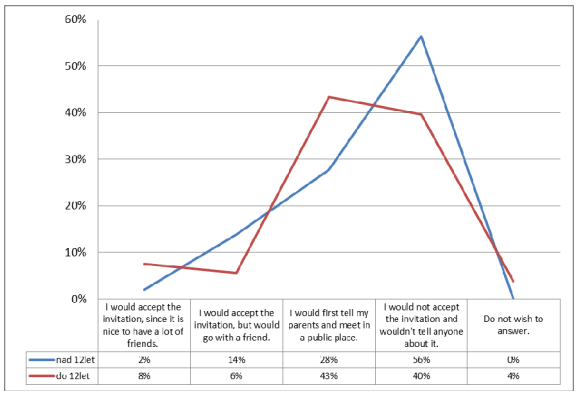


	Only my parents.	Only my best friend.	My friends when online together.	Nobody.	Do not wish to answer.
Up to 12 years of age	21,78%	15,84%	0,99%	61,39%	0,00%
Older than 12 years	64,15%	3,77%	0,00%	26,42%	5,66%

Graph 1: Who would you entrust with your passwords?

In addition to their attitude toward safeguarding their privacy while online, we were also interested in how the children and teenagers establish new contacts and create friendships online, and how they would react when faced with an invitation from an online acquaintance to meet in real life. The amount of friendships created and maintained online undoubtedly increased greatly with the wide popularity of social networks. The research results displayed that children in general do not discuss the difficulties they encounter online. More than half of teenagers included in the research stated that they would not meet in real life with a person they meet online; however, they would not tell anybody about the invitation to meet. The fact that almost 8 percent of children would accept the invitation to meet should raise concern. Furthermore, almost half of teenagers stated that in case their online "friend" would ask them to keep a secret, they would do so without additional questions.

Overall, the research displayed a concerning lack of awareness regarding the dangers of cybercrime, cyber-bullying in particular, since 46 percent of teenagers stated they are not particularly concerned regarding this issue and would simply ignore it if encountered with a case of cyber-bullying. The results of the analysis thus pointed out certain areas of concern where additional education of children and teenagers is needed in order to avoid further problems. Children and teenagers should be aware of the positive and negative points of the use of internet, and should be taught how to continue their self-education regarding cybercrime and internet safety.



	I would accept the invitation, since it is nice to have a lot of friends.	I would accept the invitation, but would go with a friend.	I would first tell my parents and meet in a public place.	I would not accept the invitation and wouldn't tell anyone about it.	Do not wish to answer.
nad 12let	2%	14%	28%	56%	0%
do 12let	8%	6%	43%	40%	4%

Graph 2: What would you do if a person you met online would ask you to meet in person?

4 PROJECT DESCRIPTION

In order to ensure the use of effective and modern learning methods, we use state-of-the-art LMS application eCampus to create hypermedia e-learning content and deliver it through blended learning approach to approximately 10 percent of Slovenian primary schools.

LMS system is a web application designed for all platforms and devices, including tablets and smartphones.

The primary objective of the project is the development and implementation of educational modules for the field of cybercrime related to children and teenagers, which are designed for specific target groups, namely children, teenagers, parents and teachers.

The educational modules include the field of information system security, cybercrime victim protection and support, online safety, online activities and communication via mobile technology, etc. and strive to reach active involvement of all participants. Special attention is paid to the youngest target group (3rd and 4th grade) by designing learning content specifically for their level of understanding and even providing different cases for boys and girls.

The objectives and methods of the project include:

- overview and analysis of critical areas of cybercrime related to children and teenagers,
- development of e-learning materials to be included in the educational modules,
- development of blended learning methods (combination of e-learning and face-to-face learning) in order to achieve active involvement of participants,

- implementation of educational blended learning modules in selected schools, and
- evaluation of results.

To achieve the planned objectives, various expert areas are incorporated in the project: educational, technological, psychological, sociological, and legal. Consequently, the experts of different fields are involved in order to ensure an all-inclusive and cohesive content, and implementation of an interdisciplinary approach. All experts involved in the project have long-standing experience and expertise in their particular field.

A project management methodology is used in all phases of the project to ensure the control over all project activities and ensure stable use of resources. Furthermore, the inclusion of relevant parties in the project team or as external evaluators aids the implementation, as well as continuance, of the project objectives after the formal end of the project.

The target group includes selected groups from different primary schools - their pupils, parents and teachers. The effect of the performed educational activities will be evaluated through comparison of a questionnaire/exam prior and following the course coupled with a questionnaire/exam one month after the course that will be performed also on a control group.

Since 10 percent of Slovenian primary schools participate in the project, the critical mass is reached in order to create multiplying effects and lead to the long-term goal of the project, which is to continue the educational activities culminating in the implementation of the developed modules in the obligatory educational curriculum.

Furthermore, the project includes the development and dissemination of best practices in the field of protection and support of cybercrime victims in relation with these types of cybercrime.

5 CONCLUSIONS

The presented preliminary research, coupled with research performed with parents and educators displayed a need for additional education of all target groups regarding the dangers of cybercrime and the importance of information safety. In order to successfully address the issue of cybercrime, it is important to implement successful preventive techniques in all target groups; therefore, continuous education undoubtedly plays and important role in raising the awareness of all users and encourages them to implement the preventive techniques in everyday life.

In order to evaluate the effects of the educational module implementation, an evaluation of the effects will be performed following the conclusion of each educational module. Through the implementation of the educational modules that will target our youngest internet users, and consequent recommendations for further actions, we will be making the next step toward the creation of an information security culture.

REFERENCES

Čehovin, G., 2010. Otroška pornografija na internetu. FDV, Ljubljana.

Gartner, Inc., 2013. Gartner Says Worldwide PC, Tablet and Mobile Phone Combined Shipments to Reach 2.4 Billion Units in 2013. Available online: http://www.gartner.com/newsroom/id/2408515 (26.08.2013).

Gregorič, U., 2010. Socialni inženiring v spletnih socialnih omrežjih. Available online: http:// www.fvv.uni-mb.si/dv2010/zbornik/informacijska _varnost/gregoric.pdf (14.04.2013).

Dimc, M. Dobovšek, B., 2012. Kriminaliteta v informacijski družbi. Univerza v Mariboru, Fakulteta za varnostne vede, Ljubljana.

Maurel, L., 2009. L'identité numérique, bientôt saisie par la loi ?. Available online: http:// scinfolex.wordpress.com/2009/07/17/lidentite-numerique-bientot-saisie-par-la-loi/ (26-08-2013).

Ólafsson, K., Livingstone, S., 2013. Children's Use Of Online Technologies. The London school of Economics and political science. Available online: http://eprints.lse.ac.uk/50228/ (26.08.2013).

Rosen, C., 02007. The new Atlantis, A Journal of tehnology&society. Virtual Friendship and the New Narcissism. Available online: http:// www.thenewatlantis.com/publications/virtual-friendship-and-the-new-narcissism (6.april.2013).

Safe.si, 2013. ABC varnosti in zasebnosti na mobilnih napravah. Available online: https://www. varninainternetu.si/content/uploads/2013/01/Varnost-in-zasebnost-na-mobilnih-napravah.pdf (31. 3. 2013).

Shea, V., 1994. Netiquette. Albion Books, San Rafael, CA.

Spletno oko, 2013. Starši in otroci, zavarujte se pred zlorabami na internetu. Available online: https://www.spletnooko.si/r/9/75/Novice/%20Starsi_in _otroci_zavarujte_se_pred_zlorabami_na_internetu/ (27.8.2013).

Data Model and Smartphone App in an Observational Research Social Network

Cristina Mesquita[1] and Rui Pedro Lopes[1,2]
[1]*Polytechnic Institute of Bragança, Campus St. Apolónia, Bragança, Portugal*
[2]*IEETA, University of Aveiro, Aveiro, Portugal*
{*cmmgp, rlopes*}*@ipb.pt*

Keywords: Observational Research, Children Observation, Smartphone App.

Abstract: Children observation is a time consuming and complex process that involves time and knowledge. However, when performed adequately, it is fundamental for both child and professionals development, allowing a sound basis for reflection and action around learning experiences and teaching environments. Following specific guidelines and programs can help the teachers reduce inherent subjectivity. The EEL/DQP was developed to evaluate and improve quality in preschool education. This program is enhanced by using a rigorous observation process of children and adults. In this paper we propose an unified data model to structure the information from the thirteen steps of the EEL/DQP. Moreover, we also describe a smartphone application to assist the observer, recording notes, sound, photos and video.

1 INTRODUCTION

Children observation is a time consuming and complex process that involves time and knowledge. However, it provides a valuable tool to support both child and professionals development, allowing a sound basis for reflection and action around learning experiences and teaching environments (Belfield et al., 2006; Schweinhart and Weikart, 1997).

There are several programs in which observations are used to assess children physical, emotional, social, and intellectual development. Well-planned observations can give us important information about specific areas, such as social interaction, learning experiences, space management and creations, and others. Observation can also help to better understand how different areas of development are interrelated, as well as helping recognizing what behaviors are typical of various age groups. In turn, this understanding will help the teacher to improve as a person and as a professional.

Specific and well-known guidelines allows the teacher to be more focused in the observation and to make more objective interpretations. Some initiatives and projects that provide such guidelines have been implemented, such as Effective Early Learning (EEL) (Bertram and Pascal, 2004), known in Portugal under the designation *Desenvolvendo Qualidade em Parceria* – DQP (Bertram and Pascal, 2009). These initiatives define a systematic and rigorous four stage process of Evaluation and Development Framework (Pascal et al., 1997).

This four staged framework includes the Evaluation of Quality of early learning, followed by Action Planning. In this step, priorities are identified and an action plan is generated. The Development phase follows, during which the quality improvement action plan is implemented and, finally, the Reflection phase, focusing on reviewing the impact of the former.

In this paper we propose using a smartphone based application to record the observation materials directly in the device. The application is already implemented and fully working, according to the following sections. This allows recording videos, photos and sounds, with the possibility of exporting to a specific social network for observational research (Lopes and Mesquita-Pires, 2014). This online, web-based, regulated community, provides a central and secure repository of observation data, results and annotations that, together with the smartphone application, mediates communication between participants and maintains a memory of observational data.

The next section describes the observation process as described in the EEL/DQP initiative. It proceeds with the design of the smartphone application, namely the user interface, data model and user input approach, in section 3. Section 4 closes the paper with some conclusions.

2 EVALUATION PROCESS IN EEL/DQP

The evaluation of quality of early learning in the scope of EEL/DQP requires obtaining a considerable amount of data through several techniques, including detailed observations of children and adults, interviewing parents, practitioners and children, documentary analysis and others. This complex and somewhat subjective process requires well-trained teachers and researchers. In particular, the EEL/DQP initiative defines a four phase/thirteen steps procedure, comprising:

1. Evaluation
   1.1. Initial preparation
   1.2. Initial data gathering
   1.3. Interviews
   1.4. Child Tracking Observation Schedule
   1.5. Child Involvement Scale
   1.6. Adult Engagement Scale
   1.7. Development of evaluation report
2. Action planning
   2.1. Development of the action plan
3. Development
   3.1. Document and support the implementation of the action plan
   3.2. Child Involvement Scale, following the same procedure from 1.5
   3.3. Adult Engagement Scale, following the same procedure from 1.6
4. Reflection
   4.1. Reflection on the impact of the plan in the future
   4.2. Final report

Data is gathered and systematically organized in research portfolios, that will be used in a cyclic process of thinking-do-thinking to research and create change (Mesquita-Pires, 2012). This process is enhanced by the utilization of observation techniques which measure the effectiveness of the learning and teaching processes, such as the Child Tracking Observation Schedule (step 1.4), to gain a snapshot of the childs day and providing information of learning experiences (Bertram and Pascal, 2006), the Child Involvement Scale (step 1.5), an observation technique which measures the level of a childs involvement in an activity, the Adult Engagement Scale (step 1.6), to evaluate the interaction between the practitioner and the child (Laevers, 1994).

2.1 EEL/DQP Procedure

The EEL/DQP overall procedure follows the four phases described above. It starts by an initial orientation of the work to be performed, in which all the process is prepared and all the participants informed in detail. Initial data gathering follows, where the institution is characterized, including the interior and exterior spaces, its education philosophy, the different learning activities, and others.

The third step includes performing interviews with the dean, staff (about 50 %), children (20%), and parents (20%). It is very important that all the stakeholders are well informed and have a full understanding of the process. The teacher records the interviews and take notes of key phrases, to support the written report. In the end, access to interview transcripts must be given to participants.

2.1.1 Child Tracking Observation Schedule

The fourth step requires an observation process, using the Child Tracking Observation Schedule (CTOS), with the main purpose of understanding the child's daily routine. This technique gives information about the learning experiences, the level of choice, child involvement, the group organization and interaction with adults.

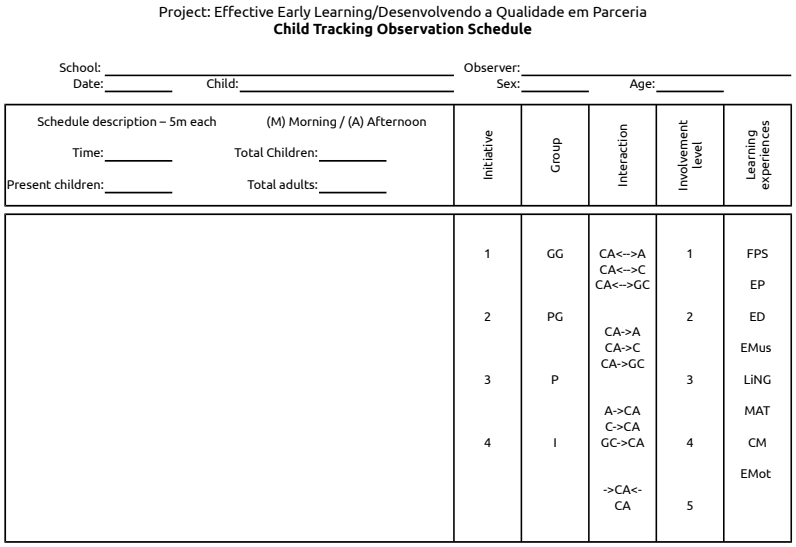
In broad terms, this step starts by choosing a set of children, considering gender, ethnicity, age, and others. Each child is observed two sets of two times, with each session in a different day. In other words, each child is observed four times, where each observation should not last longer than 5 minutes. The details are registered in a specific form (Figure 1).

2.1.2 Child Involvement Scale

The Child Involvement Scale (CIS) seeks to understand the learning outcomes and the underlying processes. Essentially, it gathers information about the participation in activities and projects, thus giving indicators of concentration and motivation as well as of satisfaction.

High levels of motivation reveals that the children are truly interested and driven to engage in the activity. The impressions and experiences obtained by a motivated child are very strong, assuming an influence that will not occur otherwise. High level of involvement also results in a high level of satisfaction. Children spontaneously take initiatives to get into this particular state. This is also achieved through the interest that the children have to discover or explore.

The concept of involvement refers to the human quality (visible in children and adults) characterized


Project: Effective Early Learning/Desenvolvendo a Qualidade em Parceria
Child Tracking Observation Schedule

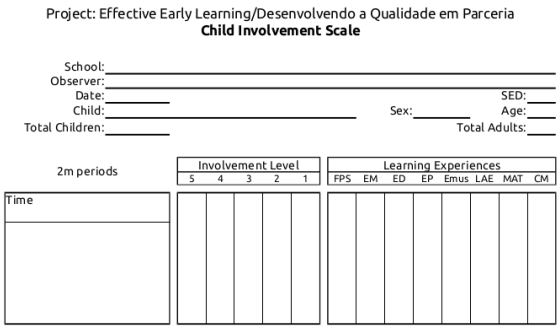
School: _____ Observer: _____
Date: _____ Child: _____ Sex: _____ Age: _____

Schedule description – 5m each (M) Morning / (A) Afternoon Time: _____ Total Children: _____ Present children: _____ Total adults: _____	Initiative	Group	Interaction	Involvement level	Learning experiences
	1	GG	CA<-->A CA<-->C CA<-->GC	1	FPS EP
	2	PG	CA->A CA->C CA->GC	2	ED EMus
	3	P	A->CA C->CA	3	LiNG MAT
	4	I	GC->CA	4	CM EMot
			->CA<- CA	5	

Figure 1: Child Tracking Observation Schedule.

by intense involvement in activities, being considered as a necessary condition to carry out a deeper level of learning and development (Laevers, 2005). Involvement contributes for a deeper way to learn. When lacking, chances are that the children development will stagnate, and all the actors in the education process should do everything in order to create an environment in which children can engage in a wide variety of activities.

Under the EEL/DQP this step requires to select 50% of the 3, 4 and 5 years old up to a maximum of 12 children. Each child is observed three times, with each session in a different day. In other words, each child is observed in a total of six times, where each observation should not last longer than 2 minutes. The details are registered in a specific form (Figure 2).


Project: Effective Early Learning/Desenvolvendo a Qualidade em Parceria
Child Involvement Scale

School: _____
Observer: _____
Date: _____ SED: ___
Child: _____ Sex: _____ Age: ___
Total Children: _____ Total Adults: ___

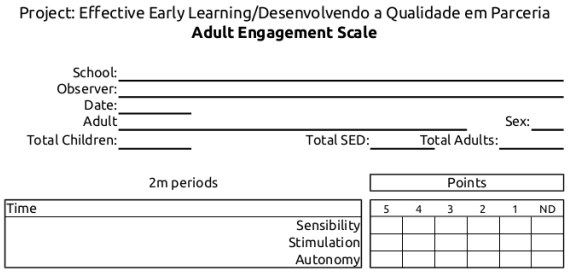
2m periods	Involvement Level 5 4 3 2 1	Learning Experiences FPS EM ED EP Emus LAE MAT CM
Time		

Figure 2: Child Involvement Scale.

2.1.3 Adult Engagement Scale

The Adult Engagement Scale (AES) evaluates the interaction between the practitioner and the child (Laevers, 1994). It targets the effectiveness of the teaching-learning process through observation of adult-child interaction. The quality of the adult's intervention is a critical factor for the child's knowledge building.

Up to a maximum of 5 adults should be observed, paying special attention to the sensibility, stimulation and autonomy categories. Each adult is observed twice, with each session in a different day. Each session should include five observations of 2 minutes each, in a total of 10 minutes. In other words, each adult is observed four times, 40 minutes total. The details are registered in a specific form (Figure 3).


Project: Effective Early Learning/Desenvolvendo a Qualidade em Parceria
Adult Engagement Scale

School: _____
Observer: _____
Date: _____
Adult _____ Sex: ___
Total Children: _____ Total SED: _____ Total Adults: _____

2m periods	Points 5 4 3 2 1 ND
Time	
Sensibility	
Stimulation	
Autonomy	

Figure 3: Adult Engagement Scale.

The observation process is translated into an evaluation report, to be used in the following (2-4) phases.

2.2 Challenges

The EEL/DQP process is heavily structured around paper based forms and reports. In this approach, many details are usually over summarized or even lost. This also prevents broader studies from being made, such as inter-school or inter-country assess-

ment, unless the notes and observation portfolios are physically shared.

In this paper we propose the design and development of an online, web-based, regulated community, providing a central and secure repository of observation data, results and annotations. This service allows each teacher to manage each observation process, sharing information with authorized colleagues and improving their experience and knowledge. It is complemented with the tools to generate PDF reports and spreadsheet either for individual as for collective studies.

We also propose using a smartphone based application to record the observation materials directly in the device. This allows storing a portfolio of observations in a single device, as well as recording videos, photos and sounds together with written annotations.

3 SMARTPHONE APPLICATION

Smartphones present a set of restrictions that are necessary to consider when developing applications for these kind of devices. Although processing power, battery and memory are potentially limiting issues, the most important aspect to consider is, probably, the display size. It is used both to present information to the user and to collect input from him. This characteristic influences the user interface, designed around sequences of pages, switched (or flipped) according to the context.

In the previous section we presented an overview of the process for quality enhancement in preschool education, comprising several paper-based forms and specific observation techniques. The form occupies all the available paper space, allowing the observer to immediately see the whole picture. If the data grows, it is necessary to move to an additional sheet of paper. Moreover, each observation is registered in an individual sheet, being necessary to organize and archive all the sheets in a portfolio.

Another limitation of the paper based form is that it does not allow to record photos, sounds or other media, requiring additional tools, such as video-camera or digital sound recorder. The process also requires that the observer pay rigorous attention to the time through a stop watch, for example (Mesquita-Pires and Lopes, 2013; Mesquita and Lopes, 2013).

3.1 User Interface

The design of the user interface has to incorporate all these aspects, in a small touch screen. Each form is accessible through a tabbed panel, allowing switching between forms with a single touch (Figure 4). The main panel allows the user to switch to the Child Tracking Observation Schedule (CTOS), the Adult Engagement Scale (AES), the Child Involvement Scale (CIS), and the synchronization tool (Sync).

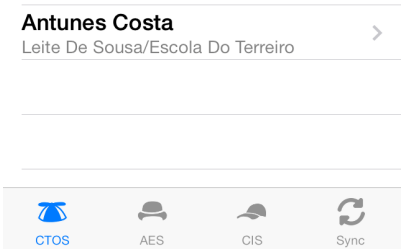


Figure 4: Tabbed panel, allowing access to other forms.

In each form, the master-detail interface paradigm is used to present a tree of contexts and associated data. This allows organizing all the portfolios and observations in a drill-down navigation model, in a comfortable to use approach (Figure 5).

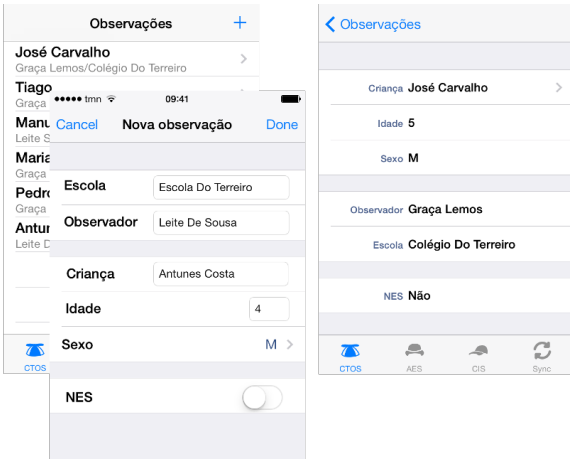


Figure 5: Observation views: listing, adding and details.

The user creates a child's portfolio by touching the '+' button, in the screen's upper right corner. A new view is shown to allow entering the observation details. Each entry in the main list shows the child's name, in bold, and, below, the observer's and institution's name. Touching an element in this list, will push a new view with the remaining details. Touching the child's name, will advance to the sessions section. The same paradigm is used throughout the application.

Just as in the previous case, the main view presents the list of sessions, identified by the date. Below the date, the number of children adults is shown (Figure 6).

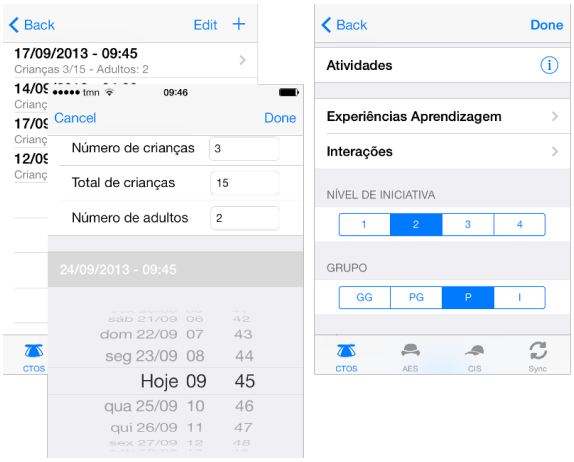


Figure 6: Session views: listing, adding and details.

By touching the '+' button, the user will be able to create new sessions. When selecting a session from the list, the view is replaced by the session details, to allow entering the related information. The initiative level, the predominant group and the involvement level is entered by selecting the corresponding button. The interactions and learning experiences will present a new view to the user where he will be able to select from a list of choices.

3.2 Data Model

As mentioned above, the observation process in the EEL/DQP project is implemented in three procedures: the Child Tracking Observation Schedule, the Child Involvement Scale and the Adult Engagement Scale. Each observation detail are registered in paper based forms (Figure 1, 2 and 3).

The CTOS form starts by identifying the institution and the observer name, the date, time, and the child's name, sex and age. In addition it also records the number of children and adults present during the session. It also registers the child's level of initiative (1 to 4), learning experiences, predominant group, involvement (1 to 5) and interaction, either verbal or non-verbal (Table 1).

There are four levels of child's initiative:

Level 1. No possibility of choice is given to the child performing the activity

Level 2. A limited number of possibilities of choice is offered to the child

Level 3. Some activities are not allowed to be chosen by the child

Level 4. The child has total freedom of choice

Table 1: Child Tracking Observation Schedule interaction.

CA ←→ A	Balanced interaction between target child and adult
CA ←→ C	Balanced interaction between target child and child
CA → A	Target child interacts with adult
CA → C	Target child interacts with child
CA → GC	Target child interacts with group of children
A → CA	Adult interacts with target child
C → CA	Child interacts with target child
→ CA ←	Target child talks to himself
CA	No interaction
GC → CA	Group of children interacts with target child
CA ←→ GC	Balanced interaction between target child and group of children

The predominant group situation is registered as follows: Large Group (GG), Small Group (PG), Pair (P), Individual (I). Learning experiences fall under: Personal and Social Education (FPS), Physical Expression (EMot), Representation and Dramatical Expression (ED), Artistic Expression (EP), Musical Expression (EMus), Oral Language and Approach to Writing (LING), Mathematics (MAT), Knowledge of the World (CM).

Children involvement follows five levels:

Level 1. Absence of activity

Level 2. Activity frequently interrupted

Level 3. Activity almost continuous

Level 4. Activity with intense moments

Level 5. Maintained intense activity

The CIS form uses the same values from the involvement level and the learning experiences, changing the way the observations are made. This allows us to simplify the data model, by reusing the same entities defined for the CTOS module.

The Adult Engagement Scale has a different header as well as focus. It defines three sub-scales, measured from 1 to 5:

Sensibility: the attention and care the adult demonstrates towards the feelings and enotional well being of the child. It also considers the honesty, empathy and affection;

Stimulation: the way as the adult fulfills the learning process and the content of the intervention;

Authonomy: the degree of freedom the adult concedes the child to experiment, judge, choose learning activities and express ideas and opinions.

Using the paper based forms as a starting point, all fields are collected and corresponding data types identified (Table 2). The fields from the CIS form is clearly a subset of the CTOS form, allowing some simplification to the data model. On the other hand, the AES form introduces different fields, that must be considered.

Table 2: Data fields and corresponding data type.

		Forms		
Data Type	**Field Name**	**CTOS**	**CIS**	**AES**
String	school	X	X	X
String	observer	X	X	X
Date	date	X	X	X
String	childName	X	X	
Number	childAge	X	X	
String	childSex	X	X	
Boolean	nes	X	X	
Number	nChildren	X		
Number	nTotalAdults	X	X	X
Number	nTotalChildren	X	X	X
Number	involvementLevel	X	X	
Number	initiativeZone	X		
String[]	learningExperiences	X	X	
String	group	X		
String[]	interaction	X		
String	adultName			X
String	adultSex			X
Number	nTotalNes			X
Number	sensibilityLevel			X
Number	stimulationLevel			X
Number	autonomyLevel			X

In complement to the information from the paper-based forms, we also want to record different media, such as text, photos, sound and videos. This data is associated to each observation, so we have to specify additional data names and types (Table 3). The fields are common to all the observations, regardless of the form structure.

Table 3: Additional fields and corresponding data type.

		Forms		
Data Type	**Field Name**	**CTOS**	**CIS**	**AES**
Date	timestamp	X	X	X
ActivityType	type	X	X	X
String	descr	X	X	X
Blob	activity	X	X	X

The resulting data is structured in two conceptual groups, one for the fields from the CTOS and CIS forms and the other for the fields specified in the AES form. Each group will have two one-to-many associations, starting with the `Observation` (`AESObservation` for AES) entity for it is the most generic. Each `Observation` has many `Sessions` (`AESSession` for AES). In turn, each `Session` records the child's behavior acting on a set of `Activities` (Figure 7).

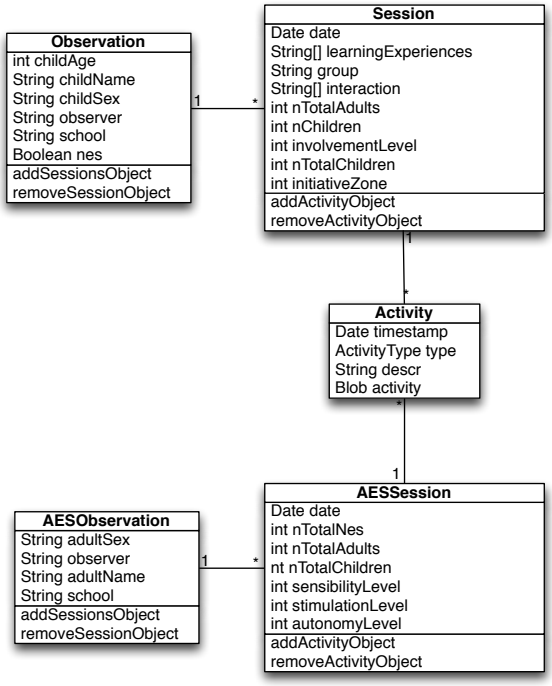


Figure 7: Entity diagram.

The `Observation` stores general information about each target child, namely, the institution's name (`school`), the observer's name (`observer`), the child's name (`childName`), sex (`childSex`) and age (`childAge`). It also registers if the target child requires Special Education Needs (`nes`). Objects of this type are the umbrella to all the information about a single target child.

The `AESObservation` stores general information about each adult, namely, the institution's name (`school`), the observer's name (`observer`), the adult's name (`adultName`) and sex (`adultSex`).

`Session` entities registers each observation. Remember that the procedure require four total sessions per child (six for CIS). Objects of this type store the summary of the observation details, such as the date the session took place (`date`), the set of learning experiences (`learningExperiences`), the predominant group (`group`), the set of interaction indicators (`interaction`), the involvement level (`involvementLevel`), the initiative zone (`initiativeZone`). Each session starts by registering the number of children in the group (`nChildren`) and in the room (`nTotalChildren`) and the number of adults (`nAdults`).

`AESession` entities registers each observation to an adult. Remember that the procedure require

four total sessions per adult. Objects of this type store the summary of the observation details, such as the date the session took place (`date`), the sensibility level (`sensibilityLevel`), the stimulation level (`stimulationLevel`) and autonomy level (`autonomyLevel`). Each session starts by registering the total number of children with special education needs in the group (`nTotalNes`), the total number of children (`nTotalChildren`) and the number of adults (`nTotalAdults`).

Each `Observation` has multiple `Sessions` which, in turn, can have several `Activity`. The latter represent all the media than can be registered, such as text notes, audio, photos and video. The remaining `AESObservation` and `AESSession` are specific of the Adult Engagement Scale. The `Activity` is the same format, since the type of media does not change.

3.3 User Input

In the paper form, each activity is described textually by the observer. We tried to maintain the same approach although allowing the user to take pictures, make videos and record sounds from the observation (Figure 8). The combination of photos, videos and text allow associating the observation to the opinion of the observer for latter assessment. It also can be used to train new observers.

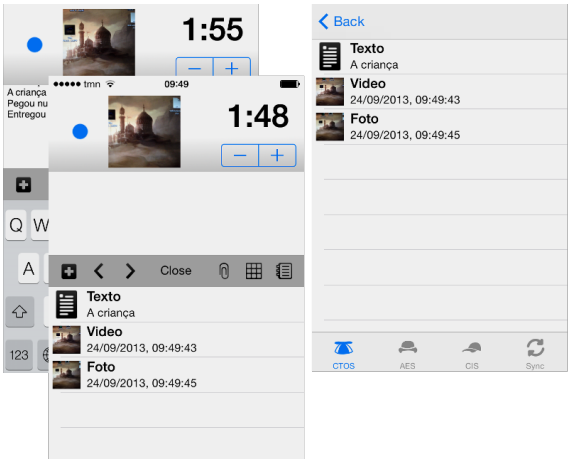


Figure 8: Activities: listing, adding and removing.

This is the view that the observer will depend the most. During the observation session, the user will write in the smartphone, take pictures and, eventually, make videos of all or pieces of the session. All these information is associated to the session that, in turn, is under an observation, thus organizing the data in an easy to browse, top-down approach – from the more general information to the more specific.

The top of the screen show a record button, to start and stop recording the video, previewed in the small window besides it. When clicking in the preview image, a photo is taken and added to the list of activities. Further to the right, there is a timer, to allow observers to keep track of the time at the same time they are paying attention to the child or adult. We added a timer (countdown) that can be set between 1 and 5 minutes. It is always visible, thus allowing for better time and observation control.

Moreover, there is also the possibility for the observer to select template phrases, thus reducing the effort of typing in the smartphone screen. The user can customize template phrases and add placeholders, to quickly type one or two words. A template phrase has the following structure:

`The child walked to the ### and grabbed the ###.`

The user can jump between placeholders (###) and type, transforming the template phrase into something like:

`The child walked to the table and grabbed the book.`

Template phrases are appended to the text area by touching it. The text is added to the activity list after pressing the '+' sign in the tool bar (Figure 9).

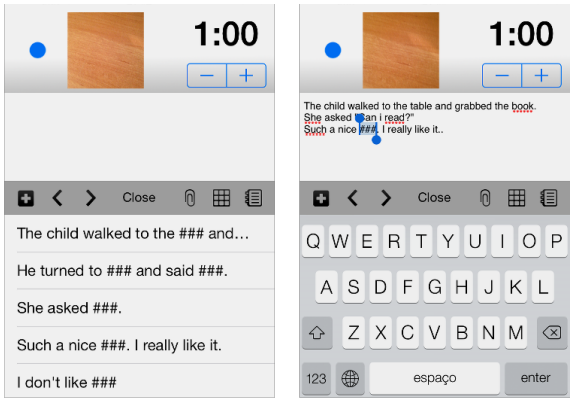


Figure 9: List and usage of template phrases.

The other tabs, in the main screen, allow the user to use different observation forms, as well as synchronizing with his area on the server. The data is stored and cyphered, so that it is protected at all times.

4 CONCLUSIONS

Observational research is a non-experimental research method in which the ongoing behavior is observed. This allows for a high degree of flexibility, because the subjects are in their natural setting. Careful and rigorous observation procedures can provide

valuable data to assess and improve quality in the processes at hand. In particular, the area of preschool education allow the professionals to improve the intentionality of their actions, the learning experiences of children and, consequently, the whole pedagogical process.

In this paper we described the adaptation of paper-based forms and methodologies to electronic format. The form fields are translated into a data model, structuring an electronic observation record. We also described a smartphone application to be constantly accessible to the professional to quickly make a note or to register a video or even photos. This data is stored in a tree structure and is later used to build observation reports.

The data can also be exported to the user personal area in a specific social network, fostering the participation of members and the maintenance of previous knowledge, that can be used for training and further studies.

REFERENCES

Belfield, C. R., Nores, M., Barnett, W. S., and Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup. *Journal of Human Resources*, 41(1):162–190.

Bertram, T. and Pascal, C. (2004). *Effective Early Learning ({EEL):} A handbook for evaluating, assuring and improving quality in settings for Three to Five Year Olds*. Amber Publishing, Birmingham.

Bertram, T. and Pascal, C. (2006). *The baby effective early learning programme: Improving quality in early childhood settings for children from birth to three years.* Birmingham: Centre for Research in Early Childhood.

Bertram, T. and Pascal, C. (2009). *Manual {DQP} - desenvolvendo a qualidade em parceria*. Ministério da Educação, Lisboa.

Laevers, F. (1994). *Adult Style Observation Schedule for Early Childhood Education ({ASOS-ECE)}*. Centre for Experiential Education, Lovaina.

Laevers, F., editor (2005). *{Sics/Zicko.} Well-being and Involvement in Care Settings. A Process-oriented Self-evaluation Instrument*. Kind & Gezin e Research Centre for Experientel Education, Lovaina.

Lopes, R. P. and Mesquita-Pires, C. (2014). Observational Research Social Network: Interaction and Security. In *6th International Conference on Computer Supported Education - CSEDU 2014*, Barcelona.

Mesquita, C. and Lopes, R. (2013). Preschool Observation Supported by Smartphone Applications. In *3rd International Conference on Wireless Communications and Mobile Computing (MIC-WCMC 2013)*, page 13, Valencia, Spain.

Mesquita-Pires, C. (2012). Children and professionals rights to participation: a case study. *European Early Childhood Education Research Journal*, 20(4):565–576.

Mesquita-Pires, C. and Lopes, R. (2013). Preschool Observation Supported by Smartphone Applications. In *International Conference on Innovation, Documentation and Teaching Technologies: New Changes in Technology and Innovation*, Valencia, Spain.

Pascal, C., Bertram, T., and Ramsden, F. (1997). The Effective Early Learning Research Project: Reflections Upon The Action During Phase 1. *Early Years*, 17(2):40–47.

Schweinhart, L. J. and Weikart, D. P. (1997). The high/scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, 12(2):117–143.

Automatic Objective Assessments of Japanese Reading Difficulty with the Operation Records on Japanese Text Presentation System
Design and Implementation

Kyota Aoki[1], Shinjiro Murayama[1] and Koji Harada[2]
[1]*Graduate school of Engineering, Utsunomiya University, Yoto, Utsunomiya, Japan*
[2]*Faculty of Education, Utsunomiya University, Mine, Utsunomiya, Japan*
kyota@is.utsunomiya-u.ac.jp, cjc42970@ams.odn.ne.jp

Keywords: Reading Difficulty, Text Presentation, Assessment.

Abstract: There are many pupils with reading difficulty in Japanese schools. The dyslexia is the disability about reading and writing texts. We use Kanji, Hiragana, Katakana characters in Japanese sentences. Based on the experiments, we show the differences between the reading activities with and without reading difficulties. This enables to make objective assessments of reading difficulty. This paper proposes the objective assessment method of reading difficulties and the design of automatic utterance detection with face images that decrease the contributions of teachers.

1 INTRODUCTION

There are many pupils with reading difficulty in Japanese elementary schools. There are many difficulties.

The big and first one is reading Japanese characters. Japanese characters are the construction of hiragana (phonetic character), katakana (another type of a phonetic character), kanji (Semantic character) and other characters. In the period of primary school, pupils learn 48 characters of hiragana, 48 characters of katakana and 1008 characters of kanji. Almost all pupils learn hiragana and katakana easily. However, the huge number of kanji is difficult to learn for some pupils in normal classes.

The next one is the difficulties about recognizing the sentence structures. In Japanese sentences, there is no spacing between words. For easing the difficulties about reading kanji characters, we can replace kanji characters with the hiragana characters. We recognize the words constructing the Japanese text in the help of kanji. There are a large number of words starting from the character of kanji. We recognize the chunk of characters that constructs a word in the complex of hiragana, katakana and kanji.

Replacing kanji characters with hiragana characters, we have the sequence of hiragana characters only. In a long sequence of hiragana, it is difficult to recognize the chunk of characters constructing a word. In our experiments, the pupils without reading difficulties like the longer high-lighted regions than the pupils with reading difficulties do. With the help of shorter high-lighted regions, some pupils with reading difficulties read more smoothly.

In an elementary school, pupils learn hiragana and katakana at first. In the first stage in elementary schools, the Japanese text-books have a space between words for the ease of understanding the structures of the sentences. However, normal Japanese texts have no space between words.

Every pupil has those 2 difficulties at first. In the long school life, they acquire the skill to conquer those difficulties. Anyway, those 2 difficulties are large barriers for reading and understanding Japanese sentences.

Every infant has no knowledge about the Japanese characters. Every pupil has a little knowledge about the huge number of kanji characters at first. Then, they learn hiragana, katakana and kanji characters in a long elementary-school life.

In Japanese elementary schools, reading difficulty means 2 years delay of reading abilities. A few of pupils with dyslexia learn in special support education schools. However, there are many pupils with reading difficulties in normal elementary

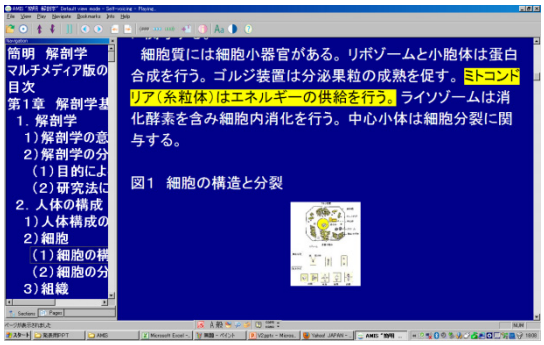


Figure 1: DAISY visual presentation example.

schools. Of course, some pupils have difficulty about remembering kanji characters. Most of the pupils remember kanji characters gradually. However, pupils with a learning disability tendency have difficulty with reading Japanese sentences in the case that they can remember the kanji characters. In the case, they may have dyslexia.

There may be many causes of the difficulties on reading Japanese texts. We do not discuss the causes. We only pay attentions to the methods for easing their difficulties. We call their difficulties "reading difficulty" in this paper.

The research about teachers shows that the pupils with ADSH tendency have difficulty about following the characters sequentially and recognizing the grammatical structures (Murayama, Aoki, 2009). Of course, there are many types of reading difficulties. There are many causes of the reading difficulties. The resulting reading difficulties show the similar symptoms. They are the difficulties about following the characters sequentially, recognizing grammatical structures and reading kanji characters.

We have developed a visual text presentation system for persons with reading difficulty in windows environments. The system records every operation of the user. With the recorded operations, we assess the difficulty of the user.

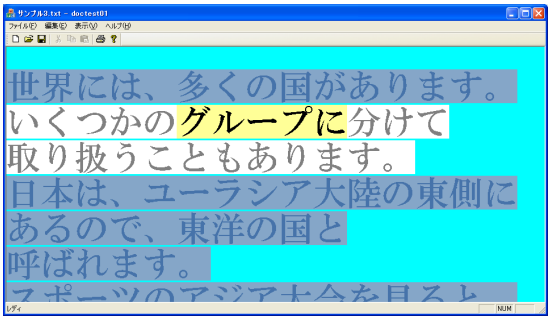


Figure 2: 1 Level annotation and 1 level mask with background coloring.

The DAISY is one implementation of digital talking book. Figure 1 shows an example of visual presentation with DAISY. It includes annotation of talking chunk of characters. It reads out the DAISY contents. It has visual text presenting functions, also (DAISY). The DAISY started for helping the people with sight disability. The DAISY is helpful for pupils with reading difficulty. However, the DAISY is a digital talking book. The content must be carefully prepared. Furthermore, there is a difficulty to make DAISY contents from electrical texts. The advanced teachers prepare the DAISY contents and use the DAISY in their classes. However, the preparations need a large amount of works. It is difficult to use emerging materials in their classes for the large preparation works. For example, it is difficult to use the morning news in the class on the same day. The interesting fresh materials can attract pupils' attentions more. The DAISY does not offer such functions. (DeMeglio, Hakkinen and Kawamura, 2002)

The Japanese text presentation system was proposed and implemented for the pupils with reading difficulties (Aoki, Murayama, 2012). The system provides the multi-level annotation as shown in figure 2. The system works well in a normal elementary school (Murayama, Aoki, 2012). The system makes the precise record of the operations. With the operational record, we can assess the reading difficulties on objective base.

This paper proposes the objective assessment method of reading difficulties with the operational record of the Japanese text presentation system, and the system decreases the work of teachers who help the pupils with reading difficulties.

First, this paper proposes the assessment method using the system operational records. Then, we discuss the problems for utilizing the assessment. Next, we propose the automatic assessment method for the pupils with reading difficulty. And last, we conclude this work.

2 OBJECTIVE ASSESSMENTS OF JAPANESE READING DIFFICULTY

With the Japanese text presentation system, the behavior of a user is recorded. From the recorded behavior, we can assess the reading difficulties about the user. This assessment is objective.

2.1 Operational Record of the Japanese Text Presentation System

The Japanese text presentation system records all the operations of a user. The Japanese text presentation system moves the high-lighted part in a text with the key-input of the user.

In using the Japanese text presentation system, the user directs the move to the next high-lighted part with a key-input. The record includes the key operations with the precise time. With the record, we can measure the time for reading the high-lighted part. Table 1 shows the example of the operational record. The 1st and the 2nd columns have the date and time. The 3rd column is the operation. The 4th column includes the high-lighted part. They are Japanese. The 5th column does the elapsed time from starting the system in 1/1000 seconds.

2.2 Assessment of Reading Difficulty

For assessing, we use the relation between the reading time and the length of the high-lighted part. There are the number of characters and the number of phonemes for measuring the length of the text. In our pre-experiments, it shows clear relations between the reading time and the number of characters. We use the number of characters for measuring the length of a text. In Japanese texts, there are kanji, hiragana and etc. As a result, there is a change of phonemes in a character. However, the number of character shows better relation to the reading time.

The Japanese text materials differ in the target age of the readers. For elder pupils, the materials include more kanji characters. A single kanji character represents a same word that is represented using many hiragana characters. Their pronunciations are same. The elder pupils read faster than the younger pupils do. As a result, there are constant relations between the number of characters and the reading time of a material.

Without reading difficulties, there is a linear relation between the reading time and the length of the high-lighted part. However, in real reading, there are many miss-operations and reading difficulties. Figure 3 shows the example of the relation between the reading time and the length of the high-lighted part. There are points on a linear function and out-lire points.

We decide the out-lire points in the reading time per character of the high-lighted part. We use a simple threshold for this process. We decide that the reading time per character without reading difficulties are between 0.1S and 0.3S. We plot the pairs of the length and the reading time of the high-lighted parts after filtering with the threshold in figure 4.

In figure 4, the pairs of the length and the reading time have the relation of linear function clearly. With the reading difficulties, the pupil needs much more reading time. As a result, the high-lighted parts where the user has difficulties for reading are plotted upper regions over the linear function.

We estimate the linear relation between the reading time and the length of the high-lighted parts from the data that are filtered with thresholds.

The data plotted in figure 4 have the relation in (1).

Table 1: Example of the recorded operations.

Date	Time	Operation	High-lighted part	Elapsed time(mS)
2013/7/4	15:42:42	RIGHT1	テレビや新聞では、	12593
2013/7/4	15:42:45	RIGHT1	天気予報が毎日伝えられています。	15375
2013/7/4	15:42:46	RIGHT1	この天気予報は、	17015
2013/7/4	15:42:49	RIGHT1	たくさんの資料にもとづいて、	19265
2013/7/4	15:42:52	RIGHT1	専門家が天気の変化を予測したものです。	22984
2013/7/4	15:42:54	RIGHT1	わたしたちは、	24250
2013/7/4	15:42:55	RIGHT1	天気予報を見て、	25765
2013/7/4	15:42:57	RIGHT1	明日の予定を立てたり、	27640
2013/7/4	15:43:00	RIGHT1	持ち物を	30281
2013/7/4	15:43:01	RIGHT1	用意したり	31250
2013/7/4	15:43:02	RIGHT1	して	33047
2013/7/4	15:44:26	LEFT3	います。	116797

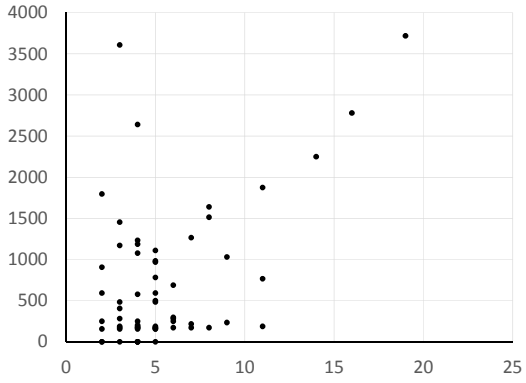


Figure 3: Relation between reading time and the length.

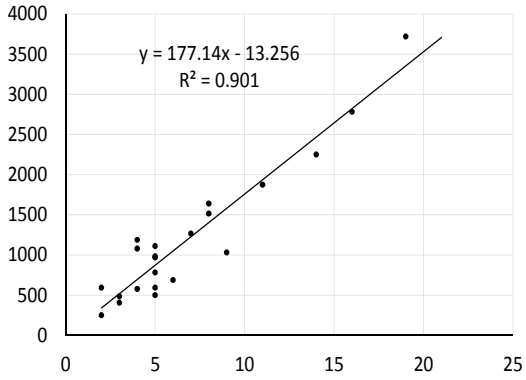


Figure 4: Relation between reading time and the length without out-lire data.

$$Y = 177.1X - 13.26 \quad (1)$$

In (1), Y is the reading time in 1/1000 second. X is the length of the high-lighted part in the number of characters.

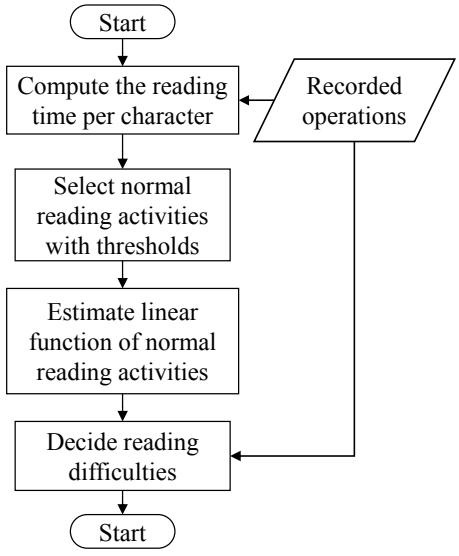


Figure 5: Assessment process of reading difficulty.

The plotted points over the normal linear function direct the reading difficulties. The corresponding part of the text shows the kinds of reading difficulties. The data under the linear function may be miss-operations. We decide that $Y \pm 1000$ is the range of normal reading activities.

The relation between the normal reading time and the length of the high-lighted part differs in the text and the school year. With large reading experiments, we can decide the normal reading activities. Figure 5 shows the process to assess the reading difficulties.

3 PROBLES ABOUT READING DIFFICULTY ASSESMENT

3.1 Teacher's Contributions

The assessment process needs large amount of teacher contributions. In reading with the Japanese text presentation system, teachers monitor the process of the readings. After that, teachers see the operational records. This assessment results an objective estimation of the reading difficulties of the user. However, there is a little difference of the teacher contributions between the assessment using the Japanese text presentation system and the classical assessment methods.

In the reading processes, a pupil may read the part that is not high-lighted. A pupil may make un-correct pronunciation. Those events make no marks in the operational record. The observing teachers guide the pupil for proper operations of the Japanese text presentation system. The teachers also record the un-correct pronunciations.

The Japanese text presentation system tries to help every pupil with reading difficulties in a normal class room. In Japanese elementary schools, there are a few pupils with reading difficulties. The teacher must make a class for the majority of normal pupils. The teachers need the day by day assessments of reading difficulties for evaluating their teaching to ease the reading difficulties of a pupil. With the present Japanese text presentation system, teachers can assess the difficulties about reading. However, the Japanese text presentation system needs many works of teachers. For enabling day by day assessments of reading difficulty, we must decrease the teacher contributions for assessing the reading difficulties.

3.2 Class Room Environments

There are many problems for utilizing the ITC technology in Japanese elementary schools (Murayama, Aoki, Morioka, 2009). The problems are listed in Table 2. For solving the problems, the proposed text presentation system treats only the electronic text. In Japan, a law forces to prepare the electronic readable texts of text books (Law). And, there are many documents accessible through the Internet. There is no paper document for an input in the proposed system.

Many pupils may remember the full text of the many times used materials as text books. Those remembered materials cannot be used for evaluating the reading performance of a pupil. The reading of the materials cannot help to enforce the reading abilities of the pupil.

In normal class rooms, many pupils use the Japanese text presentation system simultaneously. In the environments, many pupils read the text aloud. With a simple embedded microphone in PC, we have difficulties to separate each pupil's voice.

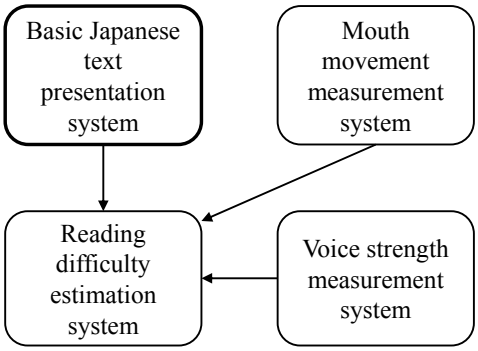


Figure 6: Japanese text presentation system with automatic assessments.

3.3 System Design

The proposed Japanese text presentation system has only 2 functions. As the first row of Table 3, we restrict the functions of the proposed system. The proposed system has some performance measurements function about pupil. This is discussed in previous section. The teachers around the pupil with reading difficulties need the objective measurements of the performance of the reading ability of the pupil. For this purpose, the proposed system provides the operation logging function. The operation logs describe the reading speed at each meaningful chunk of characters.

The proposed Japanese text presentation system enables to use one-time materials for measuring the performance of a pupil. The real-time presentation generation enables any new plain text materials at any time.

This real-time presentation generation enables to adapt the presentation for each pupil with different reading difficulties. DAISY has no function about adaptation for each pupil.

For adapting the variety of pupils' ages and disability grade, the presentation system has the function to replace the un-studied kanji characters with hiragana characters. The phonic hiragana character is first studied character. There is a little difficulty about reading hiragana.

The operations to the presentation system have the information about the user. The proposed system logs every operation with the time. This log represents the fluency of the reader.

The new system adds the video of the face images and the record of the voice to the operational record. With the video of the face image, new proposed system measure the mouth movement. With the recorded voice, the teacher may check the pronunciations afterward.

The proposed system has the features listed in Table 3. The first and the second rows are new added features. They decrease the contributions of a teacher about using the Japanese text presentation system. For wide use of the Japanese text presentation system, the system does not need large-scale contributions of teachers. The network problem is important in Japanese schools. There is a large limitation about the internet access. As a result, some cloud based implementation cannot work. The proposed system must work without the internet access.

4 AUTOMATIC ASSESSMENT SYSTEM IMPLEMENTATION

We decrease the teacher contributions with the automatic evaluation of the reading process using a microphone and a camera. The new system records the face images and the voice. The new Japanese text presentation system includes the original Japanese text presentation system. The new system includes the function to measure the mouth movement, the function to measure the voice strength and the function to make reading difficulty assessment as in figure 6.

4.1 Measurements about Voice

It is easy to measure the strength of the voice of a

user in experimental environments. With the voice of a single person, it is difficult to evaluate the precise pronunciations. However, it is easy to evaluate the strength of the voice. Figure 7 shows the wave pattern recorded with the embedded microphone in PC.

With the motions around the mouth position and the strength of the voice, the system estimates the reading activities of the user. This is much easier than to estimate the precise pronunciations.

The voice strength measurement is easy to implement. The system estimates the background noise with the measurement of continuous recorded sound. With filtered sound, we can easily estimate the voice strength of the user in the experimental environments. In figure 7, it is easy to detect the region where the voice is recorded.

In normal class room, there are many other sounds other than the voice of the user. In the environments, it is not easy to separate the voice among other voices and noises. We use the recorded voice for checking the pronunciations by the teachers.

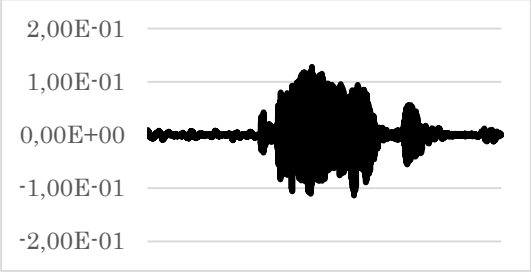


Figure 7: An example of recorded voice.

4.2 Measurements of the Motion of a Mouth

With the embedded video camera in a PC, the system records the face images of the user. Figure 8 shows an example of the processed images. The circles show the results of face and mouth detections. The figure shows the face, the eyes and the mouth clearly. The detections works well.

Figure 9 shows the measurements process of the motion of a mouth. First, face detector finds the face in an image. Then, eye detector finds the eyes. And last, the mouth detector finds the mouth in the restricted region with the detected face region and eyes' regions. As a result, mouth detector finds the mouth well. Without restriction of the search region, the mouth detector finds some number of candidates of a mouth.

Table 2: Problems about ICT usability in a special aid school in Japan.

There is a load concentration into the teacher, who is good at ICT.
There are a few educational materials for the DAISY.
They do not use the SAVE AS DAISY.
Using OCR for preparing educational materials for pupils with a learning disability tendency, the recognition errors make a large check and correct work.
There are large works for replacing difficult kanji characters with hiragana.
Using classical ICT tools as the DAISY, we need to prepare educational materials for each pupil who has a different age and a different disability.
It is difficult to evaluate the performance.

Table 3: The plan for covering the problems.

New/Old	Feature
New	Automatic operation observations.
New	Automatic assessment of reading difficulties.
Old	A collection of simple software is better than complex multi-functional software.
Old	Avoid the usage of OCR.
Old	An educational material presentation system that does not need the special material preparations.
New	An evaluation method/function for evaluating the performance of a pupil.

The system measures the motions around the mouth position. This detector is based on the face detector included in OpenCV (OpenCV). The detector has face direction compensation and face size compensations with the basic face detector. And, the eye detector estimates the eyes' positions that are not directly detected with the face detector. With the system, we can have the face and 2 eyes position continuously. With the positions both of a face and 2 eyes, we can easily estimate the mouth position relatively from the eyes' position. The mouth detector is also provided with OpenCV. In

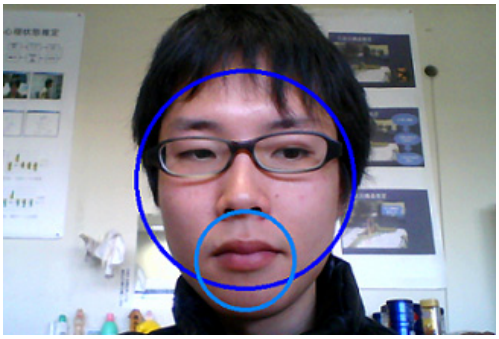


Figure 8: An example of processed images.

compensated images, the mouth detector will work well.

The face and mouth detector estimates the center and the radius of the region. With the center and the radius of the mouth's region, we decide the region of the mouth's motion detection as the rectangle circumscribing the circle describing the mouth detection. The proposed system uses this rectangle as the mouth's region.

The change of a face position is compensated with the position of the eyes. As a result, we can easily measure the mouth motions. The system measures the motion in the mouth region with the difference between adjacent 3 frames using motion template method. The method decides the mouth's region. The motion template method is also provided with OpenCV.

With this mouth motion measurement, the system estimates the reading activities of the user very precisely. With only the voice caught by a microphone, in normal classroom environment, the strength only of a voice is difficult to measure. The observation of mouth form is also difficult in normal classroom environment. However, the mouth movement only is easy enough to measure. The system measures the activity around a mouth using the motion template method.

Figure 10 shows the measured mouth's motions and utterances. The dotted line shows the utterance. The solid line shows the measured mouth's motions. The measured mouth's motion is the number of moving pixels in the mouse region. With the measured mouth's motions, we can easily estimate the utterance of a user.

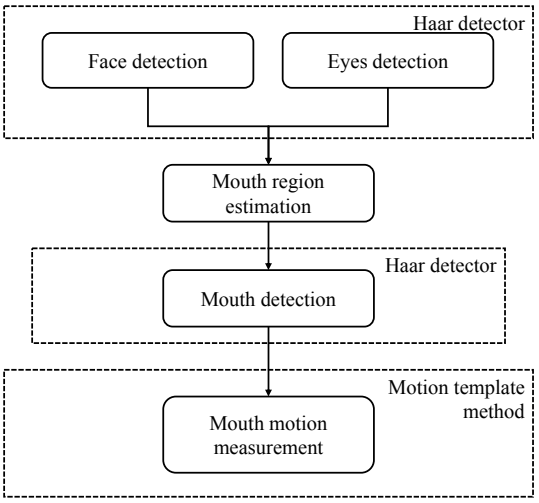


Figure 9: Mouth movement measurement.

The mouth motion measurement is robust about the changes of the environments. In class rooms, there must be proper lightings. Using the Japanese text presentation system, the user faces the display. There are no objects between the face and the display. The camera is on the top rim of the display. As a result, there is a little interception about recording the face images of the user.

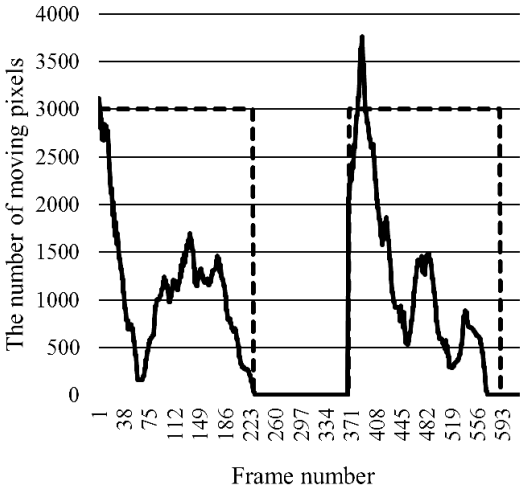


Figure 10: Measured mouth movements.

4.3 Estimation of Reading Activity

The voice of the user is easy to record using the embedded microphone in PC. It is easy to handle the recorded voice. However, in normal class rooms, there are many noises. In the environments, it is difficult to distinguish the voice of the user among many noises.

The measurement of the mouth activity is more difficult than the voice measurements, and it is time consuming. However, face image acquisition is easy and robust in the normal class room environments.

Our proposed system estimates the reading activity of the user with the video measurements of the mouth activity. When the mouth's motion is detected, the proposed system decides that the reading activity exists at the time. With this processing, the proposed system results the estimated reading activities shown in figure 11. In the figure, the dotted line shows the reading activity. The solid line does the estimated reading activity.

The error rate is 0.094 in our experiments. In a reading period, the error rate is 0.064. In a pausing period, the error rate is 0.14. With these error rate, we can check and control the operations of a user.

This implementation does not work fast enough to process in real time environments. Our implementation can process about 10 frames in 1

second. The normal video is 33 frames in 1 second. This processing speed enables to detect the utterance for teacher's checks of the automatic assessments. However, we need much more performance for assisting the user directly.

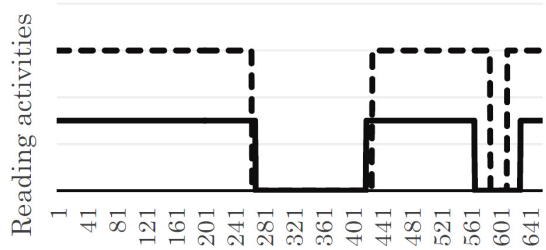


Figure 11: Estimated reading activities.

5 CONCLUSIONS

The proposed Japanese text presentation system reports the precise reading activities of the user to the teacher. The report includes not only the key operations, but also the timing of utterance with video analysis. The utterance timing is a direct description of the user's reading activities.

The proposed Japanese text presentation system with automatic assessment enables to be used simultaneously in a class room. In normal class room, a teacher controls many pupils, including ones with reading difficulties.

The new proposed system decreases the contributions of a teacher for using the Japanese text presentation system in a class. Teachers do not need to check all record of the user's reading activities. The system detects the points where the reading difficulty is. This enables easier use of the Japanese text presentation system in normal class rooms.

The much more precise record of the user's activities helps to evaluate the precise assessment of the reading difficulties with less teacher's work.

The next step of this research is a wide distribution of the software. However, this needs not only the software distribution, but also the distribution of the method to use the software effectively.

ACKNOWLEDGEMENTS

We thanks Mr. Sho Kubota for his implementation of the mouth's motion measurement. This work is supported with JSPS25330405.

REFERENCES

DeMeglio, M., Hakkinen, M., Kawamura,H., 2002. Accessible Interface Design: Adaptive Multimedia Information System (AMIS). *Computers Helping People with Special Needs, Lecture Notes in Computer Science*. Springer.

Law.http://www.bunka.go.jp/chosakuken/pdf/tokuteitosyo_fukyu_gaiyo.pdf.

Murayama, S., Aoki, K., 2010. Real Time Image Presentation System for Persons with a Learning-Disabled Tendency. *IEICE-ET, vol. 109, no. 387*, ET2009-96, pp. 25-29, 2010. IEICE.

DAISY. http://www.daisy.org/

Murayama, S., Aoki, K., Morioka, N., 2009. Image processing to make teaching aids for learning disability persons. *IEICE-108, IEICE-WIT-488*, IEICE.

Aoki, K., Murayama, S., 2012. Japanese Text Presentation System for Persons with Reading Difficulty -Design and Implementation-. *CSEDU2012, vol.1, pp. 123-128*, Porto.

Murayama, S., Aoki, K., 2012. Japanese Text Presentation System for pupils with Reading Difficulties, *ICCHP 2012, Lecture Notes in Computer Science, vol. 7382*, Computers Helping People with Special Needs, pp.507-514, Linz, Austria.

OpenCV. http://opencv.org/.

Review of Learner Modeling Using Educational Games

Mohamed Ali Khenissi[1], Fathi Essalmi[1], Mohamed Jemni[1] and Kinshuk[2]
[1]*Research Laboratory of Technologies of Information and Communication & Electrical Engineering (LaTICE), Tunis National Higher School of Engineering (ENSIT), University of Tunis, 5, Avenue Taha Hussein, B.P. 56, 1008, Tunis, Tunisia*
[2]*School of Computing & Information Systems, Athabasca University, Athabasca, Canada*
khenissi_mouhammed@yahoo.fr, fathi.essalmi@isg.rnu.tn, mohamed.jemni@fst.rnu.tn, kinshuk@ieee.org

Keywords: Educational Games, Learner Model.

Abstract: The learner model is a key component of any adaptive E-Learning system, including educational games, as it stores the information about the learner, which can then be used to provide personalized experience. Recently, there has been growing interest in creating learner models using educational games. This is due to the limitations of traditional E-Learning systems, since the learner interaction with the computer is rather restricted in such systems. Educational games, on the other hand, not only stimulate learners by increasing their motivation and engagement, but also facilitate a lot of learner interaction that could be observed to create learner models. This paper presents a survey of the field of learner modelling using educational games. In particular, it describes the main methods of learner modelling. This paper also lists several educational games that are suitable for experimentation. Compiling this information can be important to the researchers and developers working in this field, especially to new researchers.

1 INTRODUCTION

Recently, there has been growing interest in learner modelling using all types of games designed for educational purposes. Learner modelling (the process of creating learner model) requires several observations (inputs) collected from learner interaction with the computer. The learner interaction with computer is rather restricted in traditional E-Learning systems (limited to clicks, time that the learner invested in visiting the page, etc.) (Pablo et al., 2009; Stathacopoulou et al., 2007). This information does not really inform us how the learner interacted with that content. Did they read the content? Did they examine it? Did they simply go outside? (Pablo et al., 2009). This lack of information affects the accuracy of learner modelling process. On the other hand, educational games provide ample opportunities of learner interaction with the computer, which can be used to create reliable learner models. Therefore, using educational games (that represent a highly interactive content) to model the learner is a promising trend.

This paper presents a survey of the field and describes the main methods of learner modelling using educational games. This paper also lists several educational game platforms that are suitable for experimentation. This is an attempt to facilitate experimental research in this field. Compiling this information can be important to the researchers and developers working in the field of educational games and learner modelling, especially to new researchers.

The main content of the paper (sections 2 and 3) is a review of existing methods that have already been used to model the learner using educational games. To make the scope of the review more clear we define the concepts of modelling methods as set of ideas and approaches that can be presented at the conceptual level and used to achieve a specific objective. This objective is the creation of learner model.

In this review, we observed those methods that model learners in explicit way and the ones which model them in implicit way. The difference between explicit and implicit methods of learner modelling is related to the ways of extracting information about learners. An implicit method aims to extract information about learners in hidden and unobtrusive way, without endanger the high level of engagement provided by educational games. On the other hand, an explicit method aims to use a direct and obvious way of extracting information and made it overt to learners. The next section of the paper

presents identified methods that have already been used to model the learner explicitly. Section 3 presents methods that are used to generate learner models implicitly. Section 4 describes and discusses several educational games used to model learners. Lastly, conclusions are drawn and directions for future work are presented.

2 EXPLICIT METHODS

In this section, we will present identified methods that were already used to model the learner explicitly along with examples.

2.1 Evaluation of Questionnaire

Learner modelling can be performed through evaluating the learner's answers to questionnaire at the beginning, during or at the end of the educational game. It can also be done by redirecting learners to an online survey and asking learner to answer questions (Fu et al., 2009; Pourabdollahian, 2012).

The uses of questionnaire can provide direct and precise answers, but this would endanger the high level motivation provided by educational games in case of stopping learner from playing and requesting her/him to answer questions.

As examples of works that used this method, a scale named EGameFlow was used in (Fu et al., 2009) to measure learners' enjoyment of educational games. The scale contained 42 items, presented in Likert-type scales, with 1 and 7 respectively representing the lowest and highest degree to which respondents agree with the items. Similarly, in (Pourabdollahian, 2012), learners were requested to answer a questionnaire in order to evaluate their level of engagement. The questionnaire contained 21 questions based on five-point Likert scale and categorized according to five classifications (challenge, immersion, interest, purpose and control). The data collected from the survey was analyzed by both descriptive and inferential statistics for evaluating the engagement of learners.

2.2 Interpreting Body Posture and Physiological Signals

Various hardware and software equipments have also been used for modelling the learners. Behaviour of the learners can be recognized through identifying gestures, body posture and physiological signals. It is done by identifying and labelling different reactions to game events and then looking for common features of body expressions (Peters et al., 2009; Jimenez et al., 2011).

This method can provide additional information for modelling learner. However, it requires the use of additional hardware and software equipment. Furthermore, observations collected using these equipments can be interpreted in different ways and this can endanger the reliability of learner models.

(Jimenez et al., 2011) used Electro Encephalogram readings using a Brain Computer Interface (BCI) for psychometric input to measure attention. BCI offered the possibility of reading electric signals generated by neural activity in the brain, which could be used to assess the learners' attention levels. In another example (Peters et al., 2009), learner's level of attention was evaluated using two components. The first component detected user gaze behaviour based on input from a web camera. This component used either the eye direction or the head direction to establish the screen-space coordinates of where the user was looking. The second component measured attention levels using a neurophysiological recording device.

3 IMPLICIT AND UNOBTRUSIVE METHODS

Several methods are used to generate learner models implicitly. In this section, we are going to identify and discuss methods that have been used in the literature.

3.1 Translating Learner's Actions

In this method, the set of actions made by the learner during the game can be interpreted and translated into descriptive information useful to model the learner. In addition, experts or automated systems define to each set of actions an appropriate description. These descriptions are then used to model the learner (Mark, 2012; Conlan et al., 2009).

For example, the learner model in (Mark, 2012) aimed at creating a representation of learner's cognitive traits. To do that, the study established connections between the behaviour of the learner and his/her actions in the game. Specifically, the manifestations of traits that the learner may exhibit during game-play were identified. This could provide evidence of the learner's cognitive abilities.

The learner model's goal in (Conlan et al., 2009) was to generate a real-time evaluation of a learner's skills. The skill assessment engine component was responsible for translating each learner's actions

within the game into a list of probabilities that show the likelihood of each relevant skill having been acquired by the learner. The learner model determined which skill states had increased or decreased in probability.

3.2 Interpreting Interaction Traces

In this method, the game engine monitors and records interaction between the learner and the game, generating a trace of actions performed by the learner during the play sessions and tracking a lot of different parameters in the game (start game, end game, quit game, phase changes, significant variables, time, performance, tasks completed, etc.). At the end of the game play, these traces can be delivered to a group of experts or automated systems in order to identify what information can be extracted from these traces and infer a lot of useful information for modelling purposes (Stathacopoulou et al., 2007; Bouvier et al., 2013a; Bouvier et al., 2013b). For example, learner's learning style was modelled in (Stathacopoulou et al., 2007) by monitoring learner's actions over time, where each response such as a keystroke, mouse move or drag was timed and recorded. The learning environment stored all available information about what a learner was doing in a log file, recording each learner action with a time stamp. Furthermore, this work used teachers' expertise in order to select the appropriate measures of learners' observable behaviour to serve as indicators of learners' learning style.

Learners' engagement was identified in (Bouvier et al., 2013a; Bouvier et al., 2013b) by relying on their traces of interactions performed in the learning game. The study proposed an approach that consists of three stages. The first stage aims to determine the high-level engaged-behaviours. The second stage aims to characterize these engaged-behaviours by identifying the underlying chains of actions. The last stage aims to detect these chains of actions among all the actions recorded.

3.3 Interpreting Conversation

In this method, information about the learner is extracted from his/her communication with a Non Player Character (NPC). The NPC poses questions and records learner's choices during these conversations and questions answered during game dialogues (Pablo et al., 2007; Genaro et al., 2008).

For example, learner's learning style and learner's preferences were modelled in (Pablo et al., 2007) by interpreting conversations between the learner and NPC during the game play. In particular, the NPC queried the learner on his/her preferences and other questions. Depending on the learner's choices during these conversations, the game detected his/her learning style and also detected his/her preferences.

A NPC in the (Genaro et al., 2008) monitored perpetually the learner and requested her/him to answer questions. The system collected and processed the learner responses. After that, it calculated learner's level of motivation. Then it selected one rule from a set of rules taken from theory and reacted aiming to sustain or enhance the current level of motivation.

3.4 Interpreting Learner's Errors

The goal of this method is to retain the information of what the learner has learnt, and what the learner has learnt incorrectly. In addition, this method records learners' misconceptions, errors, number and type of mistake made by the learners, learner failure and success and the time taken to complete the game (Champagnat et al., 2010; Khenissi et al., 2013).

For example, learner failure and success in (Champagnat et al., 2010) are pursued. In the case of failure, the game has to be permissive with the learner. This done by punishing the learner (worse results, handicaps at the time of the following tests, less points) but not stop him from playing or reiterate the stage.

Learner model in (Khenissi et al., 2013) detected learner's deficiency in programming subject. To do that, the learner model keeps track of the learner's wrong answers when he/she arranges a set of unordered instructions in order to form a program. In particular, when the learner moves an instruction in inappropriate place, the learner model records the incorrect answer and provides him appropriate helps.

3.5 Path Follow

The goal of this method is to follow the path of learner during the game play, record learner trajectory towards a goal and then interpret and infer information about the learner. In fact, each path has a specific meaning and leads to specific information about the learner (Moreno-Ger et al., 2008; Genaro et al., 2009).

For example, (Moreno-Ger et al., 2008) described the game as state transition systems in order to assess learner activity inside the game. During the game play, actions of the learner triggered state transitions and the sequences of

actions leaded to one or many end states. The game engine kept track of transitions, checked the states that the game went through and generated reports describing them.

A virtual observer within the educational game in (Genaro et al., 2009) monitored the learner and identified his/ her learning trajectory. Learning trajectory identified was compared to different predefined learning trajectories. This comparison informed the learner model to take a decision on how to proceed.

4 EDUCATIONAL GAMES USED TO MODEL LEARNER

This section describes educational games used to model the learner. Most of these educational games include mechanism to extract specific information about learners. This is an attempt to facilitate experimental research in this field and help a new researcher in the field to start his/her own work.

4.1 Prime Climb

Prime Climb is an educational game designed to help 6th and 7th grade students practice number factorization. Prime Climb consists of series of mountains. Each mountain is divided into hexes labelled with numbers. Two players must collaborate to climb these mountains. Each player can only move to a number that does not share any common factor with the partner's number. If a wrong number is chosen, the climber falls and swings from the rope until the player select a correct number. During the game, each student has a pedagogical agent which provides support. This educational game was cited in many works for different purpose (Conati and Maclaren, 2009).

4.2 Learning Version of Pacman Game

Learning version of Pacman game (LPG) aims to motivate learners to correctly answer the questions of the programming languages. In the traditional version of Pacman game, players control Pacman through a maze, eating pac-dots. When all dots are eaten, Pacman is taken to the next stage. In addition, there are four enemies who roam the maze, trying to catch Pacman. If an enemy touches Pacman, a life is lost. Near the corners of the maze are four dots known as power pellets that provide Pacman with the temporary ability to eat the enemies. When all lives have been lost, the game ends. In the Learning version of Pacman Game, when Pacman eats a power star, the learner has to respond to a question (about the programming language) in order to continue the game having a 'reverse' role (Pacman can move freely and eat the enemy for a short period) (Khenissi et al., 2013c).

4.3 Talking Island

Talking Island is an educational Massively Multiple Online Role-Playing Game (MMORPG) designed to teach English vocabulary and conversational skills to elementary school students. It includes elements that are generally found in an MMORPG (e.g. team-work, battles, pets, and role-playing situations). The learner can practice the pronunciation thanks to voice-recognition module included in this game. In fact, this module diagnoses the correctness of the pronunciation and determines whether the learner reaches a new level or passes a quest. In addition, this game allows players to team up, communicate via voice or texts, and solve quests jointly (Hou, 2012). This educational game is accessible online at http://www.talking-island.com/

4.4 e-Adventure Educational Game Engine

E-Adventure is a game engine designed to facilitate the creation of educational games. Games delivered by e-Adventure engine are point and click adventure games. During the game, learner can learn by the interaction with objects, consultation of in-game books and conversations with other characters. The <e-Adventure> engine includes mechanisms that can monitor the learner's activities and then provides adaptation and assessment (Pablo et al., 2007). The E-Adventure platform is accessible online at http://e-adventure.e-ucm.es/

4.5 ELEKTRA Game

The ELEKTRA game is a narrative-driven adventure game. In this game, the learner has to solve several physics-oriented puzzles. A virtual character (representing the ghost of Galileo) guides, advises and encourages the learner during the game (Conlan et al., 2009). The ELEKTRA project is accessible online at http://www.elektra-project.org/

4.6 80Days Game

80Days is an adventure game used to teach

geography for a target audience of 12 to 14 year olds and follows European curricula in geography. The game story is about an alien scout called Feon which kidnaps a Boy (play character) and travels with him around the world in a spaceship to collect relevant geographical information. The player assists the alien to explore the planet and to create a report about the Earth and its geographical features. In the course of the game, the player discloses the aliens' real intention which is preparing the conquest of the earth. The player has to save the planet and the only way to do it is to draw the right conclusion from the traitorous Earth report (Kopeinik et al., 2012). The 80Days project is accessible online at http://www.eightydays.eu/

4.7 Instruction Right Place Game

Instruction Right Place Game allows learners to benefit from the drag and drop technology to construct a program (from any programming language) in an amusing way. This educational game breaks down complex programming tasks and guides learners through a series of small steps to form a program interactively (Khenissi et al., 2013a; Khenissi et al., 2013b; Khenissi et al., 2014).

4.8 Educational Games and Methods of Learner Modelling

The process of creating learner model requires several observations (Inputs) collected from the learner's interaction with the educational game.

Methods described in this paper will use these inputs in order to extract specific information about learner (Outputs). This information will be stored in the Learner Model. Most of educational games previously cited include mechanism to extract specific information about learners and then give them adaptation. Table 1 summarizes input, methods and output used by these educational games, and cited in the literature, in order to model learners.

Table 1 shows that different educational games have used different methods for creating learner model. For example, the educational game Prime Climb has used body expression of the learner, as input, and *interpret body posture and physiological signals* method in order to model learner's emotions. Indeed, manifestations of emotion are expressed immediately through the body. Hence, the best way to extract these manifestations is to interpret body posture and physiological signals of the learner using hardware and software equipment. As another example, the table shows that learner's deficiency can be modelled using *interpreting learner's errors* methods. In particular, this method tracks errors of the learner in order to interpret them.

5 CONCLUSIONS

In this paper, we focused on learner modelling using educational games. In particular, we gathered information from various works in the field and identified the most important methods used. Furthermore, several educational games used to model the learner and available for researchers are presented and discussed

The use of educational games can improve learner motivation, increase learner's desire to learn and develop positive attitudes toward many subjects. Furthermore, learner modelling through the use of educational games can create a reliable learner model that has an extremely important role in adaptive learning system.

For future work, we will look how to adapt the game experience to the individual learner using learner model. In fact, adaptation service in educational game can improve the learning process.

Table 1: Educational Games and Methods of Learner Modelling.

	Educational games	Inputs	Methods	Outputs
1	Prime Climb	Body Expression	Interpreting body posture and physiological signals	Learner's emotions
2	LPG	Questionnaire	Evaluation of questionnaire	Learner's level of knowledge
3	Talking Island	Learner behaviour	Interpreting Interaction traces	Learner activity
4	E-Adventure	Conversation of the learner	Interpreting conversation	Learner's learning style
5	ELEKTRA game	Actions of the learner	Translating learner's actions	Learner's skills
6	80Days game	Learner behaviour	Path follow	Learner's competences
7	Instruction Right Place Game	Errors of the learner	Interpreting learner's errors	Learner's deficiency in programming subject

Therefore, we will discuss methods and techniques used in this field.

REFERENCES

Bouvier, P., Lavoué, É., Sehaba, K., George, S., 2013a. Identifying Learner's Engagement in Learning Games: a Qualitative Approach based on Learner's Traces of Interaction. *Proceedings of 5th International Conference on Computer Supported Education (CSEDU 2013)*, Aachen, Germany, 6-8 May.

Bouvier, P., Sehaba, K., Lavoue, E., George, S., 2013b. Using Traces to Qualify Learner's Engagement in Game-Based Learning. *Proceedings of 13th International Conference onAdvanced Learning Technologies*, Pages 432 - 436.

Champagnat, R., Delmas, G., Augeraud, M., 2010. A storytelling model for educational games: Hero's interactive journey. *Journal International Journal of Technology Enhanced Learning archive Volume 2 Issue 1/2*, Pages 4-20.

Conati, C., Maclaren, H., 2009. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction 19 (3)*, pp. 267–303.

Conlan, O., Hampson, C., Peirce, N., Kickmeier-Rust, M., 2009. Realtime Knowledge Space Skill Assessment for Personalized Digital Educational Games. *Proceedings of the Ninth IEEE International Conference on Advanced Learning Technologies*, Pages 538-542.

Fu, F. L., Su, R. C., Yu, S. C., 2009. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education 52*, pp. 101–112.

Genaro, R. M., David, B., Sara, D., 2008. A Model of Motivation for Virtual-Worlds Avatars. *Proceeding IVA '08 Proceedings of the 8th international conference on Intelligent Virtual Agents*, Pages 535-536.

Genaro, R. M., Sara, D., Alma, R. G. G., 2009. A Model of Motivation Based on Empathy for AI- Driven Avatars in Virtual Worlds. *Proceedings of the 2009 Conference in Games and Virtual Worlds for Serious Applications*, Pages 5-11.

Hou, H. T., 2012. Exploring the behavioral patterns of learners in an educational massively multiple online role-playing game (MMORPG). *Computers & Education 58*, pp. 1225–1233.

Jimenez, C. O. S., Mesa, H. G. A., Rebolledo-Mendez, G., de Freitas, S., 2011. Classification of cognitive states of attention and relaxation using supervised learning algorithms. *Proceedings of the IEEE International Games Innovation Conference IGIC*, Pages: 31-34.

Khenissi, M. A., Essalmi, F., Jemni, M., 2013a. Presentation of a Learning Game for Programming Languages Education. *Proceedings of 13th IEEE International Conference on Advanced Learning Technologies*, July 15-18, 2013, Beijing, China, Pages 324-326.

Khenissi, M. A., Essalmi, F., Jemni, M., 2013b. Toward the personalization of learning games according to learning styles. *Proceedings of the international Conference on Electrical Engineering and Software Applications (ICEESA)*, March, 21-23, 2013 Hammamet, Tunisia, Pages.1-6.

Khenissi, M. A., Essalmi, F., Jemni, M., 2013c. A Learning version of Pacman Game. *Proceedings of the Fourth IEEE International Conference On Information and Communication Technology and Accessibility (ICTA)*, October, 24-26, 2013, Hammamet, Tunisia.

Khenissi, M. A., Essalmi, F., Jemni, M., 2014. Comparison between serious games and learning version of existing games. *Proceedings of the 6th World Conference on Educational Sciences*. 06-09 February 2014, Malta, in press.

Kopeinik, S., Nussbaumer, A., Bedek, M., Albert, D., 2012, Using CbKST for Learning Path Recommendation in Game-based Learning, *Proceedings of the 20th International Conference on Computers in Education*, November, 26-30, 2012, Singapore.

Mark, L., 2012. Identification OF Cognitive Traits Through Game Method. *Master of Science In Information Systems*. Athabasca University.

Moreno-Ger, P., Burgos, D., Sierra, J. L., Fernández-Manjón, B, 2008. Educational game design for online education. *Computers in Human Behavior*, 24, pp. 2530-2540.

Pablo, M. G., Pilar, S. T., Iván, M. O., José, L. S., Baltasar, F. M., 2007. Adaptive units of learning and educational videogames. *Journal of Interactive Media in Education*, Pages 1-15.

Pablo, M. Ger., Daniel, B., Javier, T., 2009. Digital games in e-Learning environments: current uses and emerging trends. *Simulation & Gaming 40 (5)*, Pages 669-687.

Peters, C., Asteriadis, S., Rebolledo-Mendez, G., 2009. Modelling user attention for human-agent interaction. *Proceedings of the 10th Workshop on Image Analysis for Multimedia Interactive Services*, Pages 266-269.

Pourabdollahian, B., Taisch, M., Kerga, E., 2012. Serious Games in Manufacturing Education: Evaluation of Learners' Engagement, *Proceedings of the 4th International Conference on Games and Virtual Worlds for Serious Applications*. Procedia Computer Science. Volume 15, Pages 256–265.

Stathacopoulou, R., Grigoriadou, M., Samarakou, M., Mitropoulos, D.,2007. Monitoring students' actions and using teachers' expertise in implementing and evaluating the neural network-based fuzzy diagnostic model. *Expert Systems with Applications*. Volume 32, Issue 4, May 2007, Pages 955–975.

Detecting Math-and-ICT Competence

Maria Shabanova[1] and Borislav Lazarov[2]
[1]*Institute of Mathematics, Information and Space Technologies, M. Lomonosov Northern (Arctic) Federal University, Arkhangelsk, Russia*
[2]*Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria*
m.shabanova@narfu.ru, lazarov@math.bas.bg

Keywords: Socratic Style, Dynamic Geometry Software, Van Hiele's Level, Synthetic Competence.

Abstract: Theoretically DGS (dynamic geometry software) environment allows developing students' critical thinking via discovering some properties of the figures by exploring DGS applets. Our goal is to examine how this theory works in teaching a new non-trivial geometry concept. We study which van Hiele's level of geometrical reasoning is covered by the students in two cases. The first testing group worked in Socratic style of teaching without any time limit. The second one was restricted to the regular classroom parameters and the teacher applied instructional mode of teaching. The topic is the introduction of a new locus not included in the standard curriculum but in the students' zone of proximal development. We are interested in the transition of knowledge and skills developed in paper-and-pencil context to the DGS environment. The transition we consider as an indicator of competence of synthetic type. As the outcomes of the study, some important details of incorporating DGS in teaching-learning process in secondary school were clarified.

1 INTRODUCTION

The last pan European initiatives to modernize the secondary school math education refer to so called inquiry based education (Rocard et al., 2007). Key role in the change of the paradigm play the outcomes of the SINUS and SINUS-Transfer projects held in Germany (Baptist et al., 2011) where the Dynamic Worksheets done in dynamic geometry software (DGS) environment occupied the central location in teaching-learning process.

Meanwhile several initiatives to organize more attractive and effective math education based on DGS educational environment are taking place in the Russian Federation (Shabanova et al., 2013). One of them is with cooperation of Bulgarian Institute of Mathematics and Informatics in the frame of the MITE (Methodic and Information Technologies in Education) project and it aims to design educational resources and to prepare staff for implementation of these resources in regular practice (see Note 1).

In general the goal of these movements is the requirements of the key-competences defined in (European Commission, 2004) to be covered after the compulsory education (whatever it means). We adopt the main idea of the key-competence-definition that a key-competence should be *a transferable, multifunctional package of knowledge, skills and attitudes* (ibid.) Building such kind of packages separately for any particular school subject reduces both its transferability and multifunctionality, this is why we advocate considering the competence as a synthetic concept (Lazarov, 2013). The study we present in this paper is an attempt to clarify some parameters of incorporating DGS in the regular teaching-learning practice aimed to make possible the transfer of knowledge and skills in a new context (*decontextualization*) (see Note 2). More specifically - we are interested in the level of transferability and multifunctionality of some knowledge and skills elaborated in paper-and-pencil technique into DGS environment and the potential advantages that this provides for heuristics and thus for inquiry based approach of teaching-learning.

2 METHOD AND OBJECTIVES

We consider the theory of Van Hieles level of geometrical reasoning (VanHiele, 1984) as the most relevant for describing the student's progress in building a new geometrical concept (a short description of the levels is given in the Appendix). It is important for our purpose the minimal initial van Hiele's level of reasoning to be determined as either 3 or 4. The level 3

guaranties that the student can use DGS as an educational tool but not (only) for fun. Level 4 we accept as the desired educational goal in the secondary school. Our standing point is that *a student has built the math-and-ICT component on a competence of synthetic type when (s)he is able to apply DGS for solving new non-trivial problems*. This is of course in the context of the high secondary school mathematics curriculum (grade 8 and up). Since the education in geometry is based on the traditional paper-and-pencil methodology the ability to present knowledge and skills in DGS format properly ensures their transferability and multifunctionality, i.e. the presence of a competence.

The objectives we stated were to examine the existence or lack of transfer of KSA (which indicates a kind of competence) and to clarify the time needed for elaborating a new concept applying inquiry-based style of teaching-learning. We investigate the reasoning of a new which locus, is close to the basic core of the curriculum but also requires the transfer of the KSA in a non-trivial way. We adopt the following indicators of detecting such kind of competence:

I1. Proper idea about the locus and design of a proper DGS instrument.

I2. The construction is explored and the conjecture is supported by examples; the statement is stated and proved.

I3. Ability for independent work and the level of the teacher's intervention.

We integrate the DGS interface with the common geometrical language of the van Hiele's level 3, forming a kind of metalanguage. It allows more effective communication during the lesson but also more reliable observation on the indicators. For instance measuring the transfer of KSA includes the stability of a particular DGS construction. Meanwhile we keep an eye on the student's zones of actual and proximal development (Vygotsky, 1935) when we evaluate his/her progress.

Data collection was done during and after the lesson: we gathered the sheets and the files of the students; a video record of the lessons was made. However we also used the teacher's observation on each student in a longer period to clarify the van Hiele's level reached in similar matter.

3 TECHNICAL PARAMETERS

3.1 Population

Our target group was 9th grade students of an ordinary school. We organized the experiment in two groups: the first one of 14 students and the second of nine. All of them were familiar with GeoGebra (they had more than 3 years experience in applying it). The first group is a regular class; the second group is formed by students attending a math circle.

3.2 Teachers

The teachers involved in the study were among the best math teachers, experts in inquiry-based education, authors of the textbooks and having both deep math knowledge and excellent skills in GeoGebra.

3.3 Problems

We designed two problems for the study. Problem 1. Find the locus of points inside a quadrilateral with minimal sum of the distances to the lines containing its sides. Investigate the case:

a) square;
b) rhombus;
c) parallelogram.

Problem 2. Find the locus of points inside a quadrilateral with minimal sum of the distances to its sides. Investigate the case:

d) square;
e) rhombus;
f) parallelogram.

Students were expected to pass through the following steps:

step 1: the locus is drawn correctly,
step 2: the conjecture is relevant,
step 3: the conjecture is confirmed by examples,
step 4: a proof is made,
step 5 (in Problem 1): a general proof for the square, the rhombu and the parallelogram is made,
step 5 (in Problem 2): some particular shapes of the locus are examined,
step 6 (in Problem 2): a new problem is stated and solved.

3.4 Lesson Plan

Our experiment was held in a regular lesson (2 h 45 minutes) with the second group and an extended lesson without time limit with the first group (2 h 24 min). Both lessons followed the next plan

(1) Declaring the objectives of the lesson
(2) Reminding some basic facts
(3) Stating Problem 1 and the type of the activities
(4) Individual work on Problem 1
(5) Solving Problem 1 with tutoring
(6) Plenary discussion on Problem 1
(7) Elaborating basic knowledge and skills for Problem 2

(8) Stating Problem 2
(9) Individual work on Problem 2
(10) Solving Problem 2 with tutoring
(11) Plenary discussion on Problem 2
(12) Drawing conclusions

Let us note that the concept of distance between point and segment is not introduced in the curriculum. This is why the students were given in advance a homework to elaborate the locus of points at given distance from a segment. The homework includes also problems about the sum of the distances to two parallel segments. The solutions were expected to be written in traditional paper-and-pencil style.

3.5 Benchmarks

The coverage of the indicators was registered as follows.

I1. The ability to construct a dynamically stable quadrilateral; the examined segments are dynamically stable; the sum of the segments is explicitly noted.

I2. Progress in the solution of the problems about the square, rhombus and parallelogram: the student gives arguments that support the conclusions up to the rigor proof.

I3. The student works independently or (s)he manage to obtain the required results after teacher's reinforcement.

4 LESSON FLOW

The lessons started with a brief review of the homework. For us it was a check of the initial level of reasoning of the new concept: distance between point and segment. Students performed some experimental work with the ready applets. This introductory part showed that in the both groups the DGS language was properly learned.

There were advanced students from Group 1 who managed to perform the standard questions quickly. After doing this, they went further and deeper in the solution of the problems. The unlimited time for individual work did not give any advantages to the others. As a rule, the teacher was more interested in the performance of the advanced students. She performed an individual Socratic inquiry with some of them if needed. After some time the rest of the class was given ready applets (Figure 1) provided with some short instructions.

Students of the second group were given ready applets and clear instructions what to do. This approach allowed more uniform movement along the lesson. Moreover, the teacher had direct view on the transfer

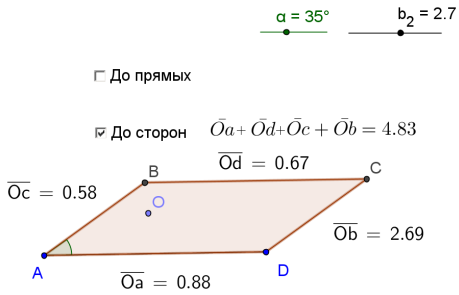


Figure 1: The main experimental applet prepared by the teacher in advance.

of the student's knowledge build in paper-and-pencil mode to the DGS environment. As soon as students performed the experimental work, she went to the proof of the results.

Very important parts of the lesson in both groups were the plenary discussions when the students presented their ideas to the class. We detected the competence level (or lack of competence) of any particular student more clearly during these discussions. The use of the metalanguage (natural language + math expressions and drawings completed with DGS-interface notation – see Note 3) allowed students to support their arguments better. We consider the ability to use properly such style of communication as existence of some synthetic competence.

5 EXAMPLES

In this section, we present our observations on 3 students. The first two are the marginal cases in Group 1: the most advanced student and the one with modest abilities. The third student is from the more homogeneous Group 2.

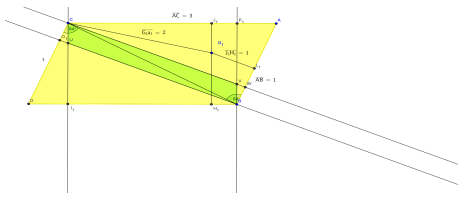


Figure 2: The only DGS construction designed by StudentI during the lesson.

StudentI (Group 1). StudentI's geometrical reasoning was classified by his teacher as van Hiele's level 4 for the most topics. StudentI often has nonstandard ideas and gives original solutions. He skipped the first problem as not interesting - he immediately guessed that the sum of the distances to two parallel lines remains the same for any point between them which answers the question in Problem 1.

StudentI designed a single DGS construction (Figure 2) but it was enough to explore the locus in the most general case. He passed the steps 1-3 in a very short time and decided he is ready. After the teacher urged him, StudentI passed step 4 - he proved the assertion without any enthusiasm (Figure 3).

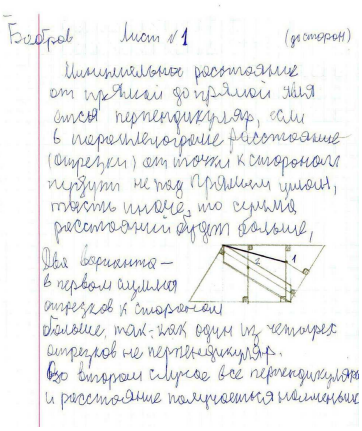


Figure 3: StudentI's P& P "proof".

StudentG (Group 1). The teacher determines StudentG's geometrical reasoning as van Hiele's level 1, but even he managed passing step 1. Further, he used the applet provided by the teacher. StudentG did not show any progress in Problem 2 after the individual work in the DGS environment. He passed step 1 in Problem 2 after some explanations given by the teacher. It seemed to be that Problem 2 is beyond the StudentG's zone of proximal development (ZPD).

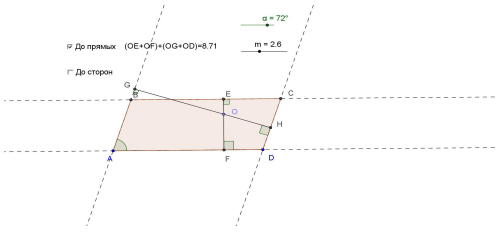


Figure 4: DGS applet provided by the teacher to StudentG with his explorations.

StudentA (Group 2). She was classified by her teacher as reasoning in Van Hiele's level 3. StudentA had some experience in a kind of research activities. She was expected to pass all steps independently. Instead of this she asked the teacher's instructions all the time. But she went further than the other students exploring some more cases. She did not prefer paper-and-pencil drawing and made her proofs using only the DGS applets for supporting her intuition.

6 RESULTS

Table 1 represents the progress made by students in solving the problems. The columns a), b), c) stand for the parts of the Problem 1; the numbers in the rows show how many students covered the corresponding step of the solution. Analogous are the columns d), e) and f).

Table 1: The progress in the solution.

Progress in Problem 1	Group 1			Group 2		
	a)	b)	c)	a)	b)	c)
step 1	14	14	14	9	9	9
step 2	8	8	8	9	9	9
step 3	13	13	12	9	9	9
step 4	11	10	6	1	1	1
step 5	1			1		
Progress in Problem 2	**Group 1**			**Group 2**		
	d)	e)	f)	d)	e)	f)
step 1	14	14	14	8	8	8
step 2	13	13	13	8	8	8
step 3	13	13	13	8	8	8
step 4	9	9	8	2	2	2
step 5	0	0	2	0	0	1
step 6	1			0		

We refer the steps 1 and 2 to the indicator I1. The success in steps 3 and 4 is related to the indicator I2. The time elapsed by the individual work or working with tutoring on the problems is connected to the indicator I3. These times were as follows:

Group 1.

Individual work on Problem 1 - 40 min.

Individual work on Problem 2 - 30 min.

Solving Problem 2 with tutoring - 17 min.

Group 2.

Solving Problem 1 with tutoring - 27 min.

Solving Problem 2 with tutoring - 17 min.

In our study we use the homework and Problem1 mainly to clarify the metalanguage in which students were going to communicate in Problem 2. The data collected clearly showed that the students were familiar with all needed auxiliary concepts when they started solving the Problem 2. The homework and Problem 1 played another important role for students according to the inductive character of the van Hiele's level - the connections between the components of the figures elaborated during solving these two preparatory problems in paper-and-pencil mode (P&P) were the necessary initial cognitive status for starting Problem 2. We give the observed van Hiele's level (the number of students who achieved the corresponding level) in table 2.

Table 2: The observed van Hiele's level.

Van Hiele's level	Group 1		Group 2	
	P&P	DGS	P&P	DGS
I	1	0	0	0
II	3	1	5	2
III	7	5	4	6
IV	3	8	0	2

7 FINAL REMARKS

We will skip advocating the obvious fact that the context of paper-and-pencil is rather different from the context of DGS environment and refer to (Marrades& Gutierrez, 2000) for the review of the research of the types of reasoning. We also agree with the findings that rediscovering nontrivial math facts in classroom practice is impossible even with most advanced students (Lazarov&Sharkova, 2013). So we are trying to evaluate a significant part of the total price that should be paid for incorporating DGS in regular teaching practice - namely the time elapsed for considerable progress in studying a new math concept. Further, after paying the price, do we get the expected quality of the educational product - the competence?

The statistics from table 1 clearly shows that the first indicator is covered by the students (with only one exception) for the most trivial case. The conclusion is that the students "speak the DGS language fluently". There is full-scale transition of the basic concepts from the paper-and-pencil context to the DGS environment. In some intuitive level students feel the matter.

The poor performance in deduction shows that the DGS abilities have little impact on the deeper understanding of the matter. This is not a big surprise if we take into account that *the deductive justification is characterized by the decontextualization of the argument used* (Marrades& Gutierrez, 2000). From our perspective the statistics indicates lack of synthetic competence of 5 students of the Group 1 and 6 of Group 2.

The StudetI's case shows the existing of KSA transferability from paper-and-pencil to DGS but the lack of it in the reverse way. This means that the step from empirical to deductive justification is not easy even for the students with relatively high math abilities.

The roots of the inquiry based approach are in the Socratic style. This style was applied in the first group after the individual work on Problem 2. The unlimited time and the Socratic style provided some opportunities for creative work to the more able students, but the tutoring instructional method applied in the second group gave better statistics.

Our initial assumption that the Problem 1 is in the zone of the actual development and Problem 2 after the homework will be in the zone of the proximal (and even actual) development of the students was too optimistic, but our misunderestimation of these very simple problems shows that the desired meaningful inquiry-based approach in math education is nothing more than a fiction for the regular classroom practice. In fact, during a regular lesson in traditional style teachers consider at least 4-5 such simple problems with rigor proofs.

On the contrary: the Socratic style is attractive and effective way for new concepts to be introduced and elaborated as extracurricular activity. During our study we observed that Socratic style teaching geometry enhanced with DGS environment does not lead to van Hiele's level 4 automatically, and in some cases even reduces students' critical thinking.

8 NOTES

Note 1. A frame of didactical technology for the Thales Theorem was among the outcomes of the 3rd meeting of MITE held in October 2006 in Oryahoviza, Bulgaria. During a workshop initiated by the second author an international team designed this frame with elements of ICT, compatible with curriculum in Bulgaria, Russia and Kazakhstan.

Note 2. We borrow the word decontextualization from (Marrades& Gutierrez, 2000) but we do not use it in the same sense.

Note 3. We use the word *metalanguage* to denote an expansion of the common math language used by the students on the DGS icons, operators etc. The general idea follows the description of the *plurilingualism* given in the Common European Framework of Reference for Languages (Language Policy Unit, Strasbourg, www.coe.int/lang-CEFR): *In different situations, a person can call flexibly upon different parts of this (communicative) competence to achieve effective communication with a particular interlocutor*.

The Van Hiele's level have specific language (Van-Hiele, 1984) which indicates the level. We observed that the students working with DGS use also in specific manner the DGS interface for each level, i.e. they 'communicate' with the computer in a manner that is specific for the Van Hiele's level reached. The communicative competence in this case is expressed via metalanguage which is specific for the corresponding Van Hiele's level. A research to clarify the metalanguage is in progress in our team and an interim report

was given at the ITA 2013 Winter Session, Dec 18-19, 2013, Sofia, Bulgaria.

REFERENCES

Baptist, P. et al. (eds.) (2011). SINUS international. Bayreuth, University of Bayreuth.

European Commission (2004). Framework for key competences in a knowledge-based society. http://ec.europa.eu/education/policies/2010/doc/basicframe.pdf (active in Nov 2013)

Rocard, M. et al. (2007). Science Education now. European Commission, Directorate-General for Research, Science, Economy and Society.

Lazarov, B. (2013) Socratic style teaching and synthetic competence building of advanced students in mathematics. DARYN, Astana, 18-19

Lazarov, B. and Sharkova, I. (2013). Two geometry illustrations of the infinitesimal method. Education and Technology. Vol 4/2013, 174-180. (in Bulgarian)

Marrades, R. and Gutierrez, A. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. Educational Studies in Mathematics, 44, 87-125.

Shabanova, M. et al. (2013) Learning mathematics with using opportunities of GeoGebra. Pero Pblsh, Moskva, p 127. (in Russian).

Van Hiele, P. M. (1984). A child's thought and geometry. In D.Fuys, D.Geddes and R.Tischler (Eds), English translation of selected writings of Dina van Hiele-Geldof and P.M. van Hiele, 237-241. Brooklyn College, N.Y. 1984.

Vygotsky, L. (1935). Dynamics of cognitive development in relation to student learning. Mental development of children in the learning process. Moscow-Leningrad, 33-52. (in Russian).

APPENDIX

We quote the definitions of the first four Van Hiele's levels as given by Michael de Villiers in *Some Reflections on the Van Hiele theory* (Invited plenary presented at the 4th Congress of teachers of mathematics of the Croatian Mathematical Society, Zagreb, 30 June - 2 July 2010).

Level 1 (Recognition). Pupils visually recognize figures by their global appearance. They recognize triangles, squares, parallelograms, and so forth by their shape, but they do not explicitly identify the properties of these figures.

Level 2 (Analysis). Pupils start analyzing the properties of figures and learn the appropriate technical terminology for describing them, but they do not interrelate figures or properties of figures.

Level 3 (Ordering). Pupils logically order the properties of figures by short chains of deductions and understand the interrelationships between figures (e.g. class inclusions).

Level 4 (Deduction). Pupils start developing longer sequences of statements and begin to understand the significance of deduction, the role of axioms, theorems and proof.

We did not detected the fifth level (Rigor) among the students of our groups.

The Life Cycle of a Cutting-edge Technology Course
A Coaching Experience on Android[*]

Maria J. Blesa, Amalia Duch, Joaquim Gabarró and María J. Serna
ALBCOM Research Group
Universitat Politècnica de Catalunya - Barcelona Tech, E-08034 Barcelona, Spain
{*mjblesa, duch, gabarro, mjserna*}*@lsi.upc.edu*

Keywords: Android, Programming, New Technologies, Google, Smart Phones, AndroidEDU, Web 3.0.

Abstract: What is the role that a university should play in the spreading of cutting-edge technologies? It is argued here that one possibility is to bring focused cutting-edge technology courses in the standard curriculum. It is contended that such courses have shorter life-spans than conventional subjects and, consequently, their implementation needs to be more dynamic. These claims are backed by discussing the *life-cycle* of an Android course running biannually from Spring 2010 to Spring 2013 at Universitat Politécnica de Catalunya. The *rise* phase of this course (which lasted two semesters) was a challenging experience that motivated students and lecturers to play a cooperative and active role in the creation of true working Android applications. The course held *stable* for two semesters while student motivation began to fall as smart phones increasingly became everyday objects. During these two phases the course was offered as *extra curricular* in the undergraduate phase. Two added factors were instrumental in the *decline* (or *fall*) phase: the availability of on-line information and the fact that the course became a requirement of a master's curriculum.

1 INTRODUCTION

Concrete IT technologies shapes our believes and behavior (Norman, 1988; Norman, 1998) and in particular, the media affect the way of teaching (Postman, 1984). From an economic point of view, technological development determines (at least in part) the role of countries in recent history. In particular the rise and fall of technological leadership deserves careful study (Nelson and Wright, 1992). In fact, the degree of technological development of a country seems to be influenced by the way in which technology has been taught (see Section B of (Nelson and Wright, 1992)). So, the importance of the role of cutting-edge technology courses is recognized not only among economists and historians but also among participants in educational meetings (Frank and Waren, 1982):

> *One of the more difficult tasks in this era of adopting curricula, is to keep a program current with technology.*

This paper analyzes the complete life-cycle (rise, stability and decline) of the course: Programming Workshop on Android Applications for Google Phones (PWAAGP from now on). This course was a cutting-edge technology course based on Android and Google phones. It originates from a synergy between lecturers from Universitat Politécnica de Catalunya and Google. Due to the very special characteristics of PWAAGP , the life of this course was short (Spring 2010-Spring 2013). Lecturers experienced such a course as an exciting (sometimes exhausting) experience in a highly volatile environment and they view its evolution as a personal challenge. The purpose of this paper is then to explain why the authors believe that such a course has necessarily a short life-cycle life and that its fall is not a failure of the organizers but a consequence of the intrinsic nature of the course. Moreover, it is claimed that many of the situations encountered in PWAAGP are common to any other IT cutting-edge technology courses.

To the best of our knowledge, there is information about the life-cycle of various teaching approaches using IT technologies as tools (Bryan, 2013). However, information about the natural life-cycle of courses on IT cutting-edge technologies is scarce.

[*]This work was supported by project 2009-SGR1137 of the Generalitat de Catalunya and grants TIN2012-37930 and TIN2007-66523 of the Spanish Government.

2 CONTEXT AND GOALS

Nowadays it is well established that universities should be a driving force for innovation. A technical university is a living and multi-headed organization and a fundamental issue to considered is how it deals with cutting-edge technologies and what is the role it should play in their standardization.

Aware of this fundamental need a team of lecturers (belonging to the ALBCOM research group of the Departament de Llenguatges i Sistemes Informàtics of the Universitat Politécnica de Catalunya decided to participate in Google's *AndroidEDU EMEA* program (AndroidEDU EMEA, 2013) in November 2009. *Google* launched a call for project proposals to universities through its *AndroidEDU EMEA* program (AndroidEDU EMEA, 2013), the aim of this program was to promote the development of teaching experiences using the *Android* operating system for mobile phones. To participate in this program, a university site had to submit a proposal for the development of an innovative system in mobile engineering that, requiring a limited number of mobile smart phones, could be addressed to a broad audience. The proposed system should also be scalable, and the proposal should promote the practical teaching of computer science issues for students.

The proposal submitted by the authors of this paper was one of the selected ones, and to facilitate its implementation *Google* donated 20 smart phones to the Departament de Llenguatges i Sistemes Informàtics of the Universitat Politécnica de Catalunya.

This is the origin of the *Programming workshop on Android applications for Google Phones* (PWAAGP). Due to the special typology of both the topic and the students, this course had a life-cycle different from regular technical courses.

This course is used as an example of a focused cutting-edge technology course with the belief that the experience obtained on this course is applicable to other similar cases. It is also claimed here that a main characteristic of this kind of courses is that they have a very short life-cycle that corresponds to the time needed by the technology to spread and become popular.

The life-cycle of this course is characterized into three shorter periods that are called *rise*, *stability* and, *decline* (or *fall*) and that follow the progression of Android from cutting-edge IT technology towards its market take-up.

3 ORIGINAL COURSE

The PWAAGP course started as an *extra curricular* subject that could be taken by students from different majors or specialties. The quota of students in the workshop was restricted by the number of phones we had. PWAAGP was a 4-ECTS[2] very practical course that required from enrolled students to have a good knowledge of Java programming language and a good general knowledge of algorithms and data structures. The course ran for ten weeks, and it was scheduled so that every week a new topic is explained through various examples and putting special emphasis in their implementation. Every week, the complexity of the issues treated was gradually increased.

Having these premises in mind, we fixed the course contents as described in the following subsection. Those contents conformed the main elementary fundamentals on which an application (or App) can be implemented. Although the contents remained almost fixed over time, that was not the case for the teaching strategy and methodology used. Conditioned by how the profile of the students was changing, and by how the technology was getting popular, we had to adapt the way those contents were taught and, consequently, they way the students were evaluated.

3.1 Contents

The first session was dedicated to the installation of the programming environment and all the development tools needed along the course. The applications and the widgets make up the highest layer in the project architecture of the Android system. The second session of the course explored the structure of these applications and presented more advanced widgets (e.g., static and dynamic lists). The topic of the third session was location. It included the embedded use of the GPS, together with the use of some hybrid web applications combining maps with other types of information. Session four introduced the SQLite, the small and powerful SQL database engine included in Android. SQLite is also used in PHP, iPhone, Skype, Mozilla Firefox, etc.).

Android inherits the *threads* directly from the Java programming language. Threads allow the execution of several tasks in parallel. This feature is especially useful when programming mobile phones, where of-

[2]ECTS states for European Credit Transfer System, which is a standard for comparing the study attainment and performance of students of higher education across the European Union and other collaborating European countries. One academic year corresponds to 60 ECTS-credits that are equivalent to 1500 – 1800 hours of study.

ten different activities are running at the same time in the device. Threads also help on doing a more efficient and user-friendly management of the applications. Threads were addressed in session five.

The subject of session six is of particular importance in Android: *Intents* and *Remote Procedure Calls*. These effective inter-process communication mechanisms increase the modularity and independence of the processes run by the applications, and facilitates the reuse and replacement of components.

In addition to its basic utility as phoning device, smartphones are also used as mobile devices to access the Internet. The seventh session was dedicated to the study of external communications by means of the WebKit browser, HTTP connections, web services, and Bluetooth. Some fundamental aspects of the graphic libraries in Android were introduced in the eighth session. In session nine, the students learned how to handle phone calls and messages while running other applications, and how to provide automatic responses to SMS. As a natural conclusion to the course, the last session was devoted to how to publish an application on the (formerly called) *Android Market*, renamed nowadays as *Google Play App Store* (Google Play App Store, 2013). Often this session was also used to have an invited talk from someone working in a mobile communications' company. That provided the students with real information, impressions and feedback from the outside world and allowed them to establish connections with the industry.

It is worth noting that this course was thought as a general post-intermediate course in programming and it was not planned for a wide audience. It is interesting to compare the contents of this course with other proposals. Consider for instance the topics covered in (Muppala, 2011). Both courses cover roughly the same topics, however in (Muppala, 2011) much more emphasis is done in the real-time aspects.

3.2 Material

This course has used a significant amount of educational and support material, including electronic devices, paper books and online material.

Google Phones. The *AndroidEDU Program* from Google awarded our course proposal with 10 phones Android Dev 1, and 10 mobile Phone Nexus One (HTC G1, 2013; Nexus One, 2013) (see Figure 1). The Android Dev Phone 1 was a variant of the HTC T-Mobile G1 designed exclusively as a tool for application developers to Google's Android OS. Both phones run on Android (1.0 Phone in Dev 1 and 2.2



(a) Android Dev Phone 1 (b) Nexus One

Figure 1: Smartphones provided by Google in 2009.

in Nexus One), have Wi-Fi (802.11 b/g), Bluetooth, quadband GSM radio and USB. The storage ability is considerably higher in the Nexus One (512MB RAM, 4 GB microSD card) than in the Dev Phone 1 (192MB RAM, 1 GB microSD card) as well as its processing power (1 GHz CPU vs. 528 MHz CPU).

Although the Android software development kit (SDK) includes an acceptably good emulator, the students found always very attractive and interesting the possibility of having real smartphones in their hands to test their applications. At the first editions of the course, the provided smartphones were of the latest technology (and therefore quite expensive) and they were not yet available to the general public. By that time, not many smartphones were available in the market. That was a major plus for the course, since the students were extremely interested in such a novel hardware and showed themselves eager for knowing how to program it using Android which, was also something to discover by that time. However, as the time passed, smartphones became more powerful and popular. That happened relatively fast. In a short time, namely three to four years, our smartphones became old and the students themselves owned better ones. Also Android established itself as an important and well-known system and everyone had heard about it at the later editions of the course. Thus, both the hardware and the software involved in the course, were gradually losing their attraction to the students.

Bibliography. At the first edition of the workshop, there were very few books on Android. The following ones were used:

- M. Murphy. *Beginning Android*. Apress Ed. 2009
- Z. Mednieks R. Rogers, J. Lombardo and G. Blake Meike. *Android Application Development*. O'Reilly Media Ed. 2009

Both are very practical, provide a good knowledge of the Java programming language, and are very focused on the implementation. They include many discussions about the code of various applications, and show how the efficiency of the programs can be im-

proved by gradually adapting them to consider the specific restrictions imposed by limited memory and battery life in mobile devices. An important feature of these two books is that the publishers offer on the Internet the source code for all programs referred in them, so it is easy to access the code, to test it, modify it and reuse it. In our opinion, they remain still two of the best books to follow for teaching Android. New editions of them exist (Murphy and Allen, 2011; Rogers et al., 2010). Since 2009, there was a burst of edited books and e-learning support material about Android on the Internet (guides, videos, etc..). Among the recently appeared references, it is worth mentioning (Haseman, 2008; Ableson et al., 2009; Mednieks et al., 2012; Burnette, 2009; Ledford, 2009; Miller, 2008) because of the quality of their contents.

Another essential online resource for this course is the Google's *Android Developer's guide* (Android Developers' Guide, 2013). In 2009, this guide was very scarce in contents and very chaotic, which made it often difficult to find specific information. Nowadays, the guide is very complete and describes basically all what can be done with Android in a very well organized way. The examples documented therein are also very appropriate.

Website and Online Forum All the contents and materials of the course, as well as projects from previous editions of the course can be found at `http://albcom.lsi.upc.edu/tpaagp/`. An online forum for students and teachers can also be found.

3.3 Evaluation

The evaluation of the students was divided into two parts: a part of continuous assessment and the evaluation of a final project. Although these two parts were always present, the way they were brought into practice had to be adapted because the profile of the students, the technology and the relation between them both, changed in the meanwhile.

Apart from introducing our students to Android, this course was aimed for enhancing them to other important abilities. For example, to work in groups, to develop a big project in a continuous and progressive way, to discuss troubles and their solutions, to brainstorm, to defend their work in English to a wider audience, to promote their work in social media platforms, etc. For working up these goals into practice, students were always asked to work in 2-3 people teams.

Continuous Assessment. At the end of each session, a programming exercise was left as homework to be solved for the next session. In order to provide the students with a general overview of the different parts conforming a non-trivial Android application, the exercise left as homework one week was, as far as possible, an extension of the one from the previous week. At every session, one team presented how did they solved the homework, and their solution was discussed among all the other students and teachers.

For evaluating this part, it was weekly taken into account how the homework was done, how did the students publicaly present their solutions to their colleagues, and also their ability to discuss and defend their solutions to the problems found.

Final Projects. The students were also asked to develop a final project. The final projects were of greater difficulty than the exercises performed along the course, and each team is assigned with a different one. Usually the final projects were proposed by the lecturers, although we exceptionally agreed sometimes on projects proposed by students. The teams had approximately two months to work on them. During that time no more pressential sessions of the workshop were scheduled.

For the evaluation of the final project, the students had to provide the programming code of the developed application, and a technical report about it. They also had to make a public oral presentation of their project in which they explained to their colleagues how they designed and programmed their application, which problems they found, how they solved them, etc. In that presentation, they also had to show that their application was properly installed and successfully running in their phones. In the first editions of the course, the students were also asked to film a video in which they promoted their application to a broader audience. Those videos are available in an dedicated channel on YouTube. All the material to be evaluated had to be in English.

We briefly explain below some of the projects undertaken to date. More information about them and all the other projects can be found at the website of the course.

Our city, Barcelona, is a metropolis with more than 1.5 million inhabitants, within a urban area of 4.5 million people. Moreover, it gets around 7 million tourists every year. That was always an inspiring scenario for us when thinking about projects to propose to the students. The proposed projects were aimed to help, by means of the mobile technology, manage some important city services as, for example, the bike rental service, car parking, car sharing, taxi finder, etc., and supporting and promoting the tourism. It is worth to highlight that most of the Apps resulting

from these projects were a big novelty by the time they were developed, e.g.:

- *gBicing.* The *Bicing* system for bike rentals has been a quite successful idea that has considerably increased the mobility within Barcelona in an ecological and sustainable way. This application allows to manage in real time the location and occupancy of the Bicing's bike stations.
- *Easy Parking.* This application helps the user to find it by means of the GPS coordinates of the car location and the use of Google Maps. When sharing the car with friends, the application gives the possibility to determine whether the car is available and an estimation of the time to retrieve it. Furthermore, this application allows to reserve the car for a period of time.
- *Guide and vote.* This application guides the user along his visit to a site and the several points-of-interest in it. Each point of interest is ranked according to the interest of its previous visitors. Once visited, the user can also vote himself.

Another type of projects were those implementing interactive games. Projects implementing some games were always welcomed by the students. They were always aimed to be played by a group of people and in a distributed way, which is quite of a novelty. E.g.:

- *Distributed Tetris.* Tetris is probably the most popular puzzle video game ever. This application allows to play Tetris in a distributed way by using radio-based communication. Not only one, but several players play the same game against the machine, and so they share the same board and have full information about the game.

4 LIFE-CYCLE

In this section, the three phases detected in the life-cycle of the PWAAGP course are described. They define the natural evolution of the course along six semesters, with a total amount of hundred student inscribed in it.

4.1 Phase 1: Rise

Covers the period from Spring 2010 to Fall 2010. Starts with the participation in the AndroidEDU EMEA program. It is offered as an *elective* subject opened to all undergraduate students. This period has its own sociological points.

- Lecturers, most of them teaching programming courses, are far from to be familiar with Android. They aim to socialize Android (open code) as an alternative to the Apple-iPhone perspective. There are very few books and the web pages are incomplete and change quickly. As information is difficult to obtain, lecturers introduce the material in formal lectures. The amount of work required to prepare a lecture is significant.
- Students don't have smart phones. They appreciate very much the opportunity to work with a real mobile phone (provided by the university). The class-group is quite heterogeneous containing computer science and telecommunications engineers. Students are highly motivated and become good friends. The relationship between computer scientists and telecommunications engineers is excellent. Some students believe that good skills in Android can help them find interesting jobs.
- Relationship between students and lecturers is very good. Students accept quite well the limited Android capabilities of the lecturers to teach some topics or to solve some specific problem. They are very collaborative and help to solve problems. The chat of the course is a fundamental communication channel.
- The evaluation is project oriented. The project is proposed at the end of the training process. As the project is "something new" it involves an important amount of extra work.

4.2 Phase 2: Stability

Covers the period from Spring 2011 to Fall 2011. PWAAGP is also offered as an elective subject.

- As web information and books about Android become more popular, regular lectures become less and less important. Students continue to be highly motivated. Slowly they become much more independent of the formal lectures and autonomous learning becomes more and more important. Formal lectures begin to disappear because students find them redundant. The role of the lecturers shift to the organizational aspects of the course.

4.3 Phase 3: Decline

Corresponds to the period from Spring 2012 to Spring 2013. Many factors contribute to the decline of this course. First of all the smart phones become a commodity. Getting information about Android is much more easy. For instance, several student associations offer Android courses. On the other hand, as it was a "successful" course, the lecturers were asked to adapt it to the standard continuous evaluation of optional

courses so as to integrate it in the near future in a master curricula. The course was lately offered as a elective course in a computer science master degree.

- Formal lectures disappear because most students find them redundant. The role of lecturers become mostly supervising. The course is reorganized as lab sessions. The session assignments are highly guided. Less free-room for creativity.
- Evaluation continues to be project oriented. The project is cut in small parts and the different parts have to be completed along the course. At the end of the course only an integration remains to be done. The number of good projects fall rather radically.

5 RESULTS

The grades of the students in each term have an important correlation with the three stages of the life-cycle of this course, as seen in Table 1.

Rise. One can see in the first two rows of Table 1 that most of the enrolled students have completed the final project with excellent results, very few of them did not work as expected and for unknown reasons some of them abandoned the course (the lecturers suspect that this was due to the significant workload required). Given these results and the feedback received from the students of these two editions of the course it can be said that they had a positive overall assessment of the course and in particular they:

- found the course contents new and interesting,
- liked to have teachers that guide the search for knowledge more than experts in the field,
- appreciated the research required by the subject,
- found the support material good and appropriate,
- appreciated the existence of weekly exercises,
- agreed with the method of assessment,
- enjoyed the final projects and found them interesting and of appropriate difficulty and,
- were convinced that the course provided useful knowledge for their future careers.

In general, the results indicate, in our opinion, that a high interest in the subject of study compensated non-experienced lecturers and/or non-sophisticated assessment mechanisms. For lecturers, the results observed exceeded by far their expectations. First, students were much more proactive and motivated than they usually are in other more conventional courses, which resulted in student projects of high quality, both in their contents and in their presentation. Second, there was an interesting change of role for the lecturers, who were acting as a learning *coaches* rather than as conventional lecturers. Third, the use of new technologies in this workshop brought new insights to classical computer science contents such as algorithmic schemes and data structures. In view of the results obtained, new technologies have shown to be a very useful motivating element for teaching that, in our opinion, is still unexploited. This experience forced us to take new positions and educational challenges. From the viewpoint of the university, this course give us the opportunity to develop university-industry relations.

Stability. The third and fourth rows of Table 1 (in sepia) correspond to the results of what was called the stability phase. In contrast with the results of the rise phase, almost all the students have excellent final projects–specially in the edition of Fall 2011, so that, it is impossible to grade any of them with honors (MH) since it was impossible to distinguish a best project among all of them. For the whole phase, only one student failed and six of them dropped-out. However, even if in Fall 2011 there is a peak in students performance, they were not as motivated as before.

For lecturers these symptoms suggested that the course workload was heavy and did not correspond to the number of credits students were granted. At this point, lecturers decided to change the evaluation method and the course entered in its decline phase.

Decline. Instead of grading a big final project at the end of the term, lecturers decided to grade, the continuous development of a single project (the same for all students) that was appropriately divided in modules to be delivered weekly. The notes of the students when this evaluation method was just established corresponds to Row 5 in Table 1 (the first in brown). There it can be seen clearly that the notes are divided into two groups: the students that performed the minimum effort just to pass the course and the ones that were really motivated and wanted to obtain the maximum grade. And this marked the end of the course since in Fall 2012 it was not offered and in Spring 2013 it became an optional master course, only 5 students enrolled and their interest was very limited, as shown in the last row of Table 1.

6 LESSONS LEARNED

What has been discussed so far is a six semester experience with a course on Android as an example of

Table 1: Statistics of the PWAAGP course. In each column it is reported the number of students that obtained the corresponding grade (going from sufficient (S) to honors (MH).

		Enrolled	NP	S	A	N	E	MH
Spring 09/10	2009-2	14	1	0	2		8	3
Fall 10/11	2010-1	20	6	0	1	1	12	0
Spring 10/11	2010-2	21	3	0	0	7	8	3
Fall 11/12	2011-1	21	3	1	0	0	17	0
Spring 11/12	2011-2	21	1	0	8	0	8	4
Fall 12/13	2012-1	0						
Spring 12/13	2012-2	5	0	3	0	2	0	0

teaching cutting-edge technologies. Although this experience was influenced by particular circumstances depending on the prevailing university context, timings, students, and work team, one can extract the following lessons:

Lesson 1. Technical Universities Should Play a Role in the Dissemination of Cutting-edge Technologies. Nowadays it is well-acknowledged that universities are expected to collaborate with leading firms in technology-transfer projects and public funds are usually available to this end. Cutting-edge technologies courses are a great opportunity to achieve this kind of collaboration and to promote the dissemination of new technologies. This collaboration is advantageous for universities because they can get free access to equipment, motivation for its students and lecturers, social visibility, and so on. Leading companies also benefit from cutting edge technology courses, since they get, among other things, feedback on products from researchers and good students, new and fresh ideas and proposals (from course prototypes), adept future engineers and public visibility.

Lesson 2. Universities Should Provide Flexible Structures to Include Short Life-cycle Courses. Our experience shows that the benefits of early adoption of technology are substantial but returns diminish rather quickly. Many universities (e.g., ours) have convoluted procedures for curriculum change—elaborate proposals, fixed and infrequent submission deadlines, several committees. It will be difficult to bring cutting-edge technology in the classroom unless some sort of ad-hoc technology-adoption policy that makes short life-cycle subjects easier to approve, implement and update is put in place.

Lesson 3. Lecturers Should Watch for Challenging Teaching Opportunities. From Lessons 1 and 2, it is a natural consequence that the right way to introduce cutting-edge technologies courses maintaining their dynamism is bottom-up. Therefore it corresponds to lecturers and leading technological firms to look for collaboration opportunities with the goal of disseminating promising technologies.

Lesson 4. Cutting-edge Technology Courses Are Extremely Dynamic. In Spring 2010, Android was emerging as a new operating system and at that time it could be considered as a cutting-edge technology but after three short years it was no longer the case. Now, Android is, a standard operating system (in fact, the most used) for smartphones, tablets, smart TVs . This standardization was accompanied by several factors: (a) The phones received quickly became obsolete and, soon, students began to own their own phones; (b) courses on Android mushroomed in internet, companies and, academies; (c) Abundant information became available in several formats (books, web-pages, web-sites, etc.); (d) Several companies developing mobile applications emerged; and (e), Being proficient in Android passed from being a plus to being a requirement.

Lesson 5. A Life-cycle Appears to Exist. As shown in this paper, in only six editions the course completed its life-cycle. One can see that this life-cycle matched the evolution of Android from cutting-edge to a market standard. In fact, this case suggests that the life-cycle of college courses that introduce a cutting-edge technology is linked to the progression of the technology towards its market take-up. The claim is that cutting-edge technology courses should form part of college curricula, but always being aware that their life-cycle is short and dictated by the standardization of the technology.

Lesson 6. Motivation Changes. During the *rise* phase of this course, Google was keen on the dissemination of Android among college students, the students were highly motivated and well-disposed to learn and committed substantial amount of time to this purpose. Lecturers were motivated also by, both, the challenges of learning and teaching Android and the commitment of students. This situation was kept for a couple of semesters until students ceased to see

this knowledge as an advantage (for instance for getting a job). and became less inclined to follow the guide of teachers who couldn't keep up with the pace of that technology. From the lecturers point of view, updating course contents from one semester to the next, involved a substantial amount of work that had a modest lasting value, minor recognition from the university and, at the end, no recognition from students.

7 CONCLUSIONS AND OPEN QUESTIONS

An experience teaching a course on *Android applications for Google Phone*s is presented. The course originated from the *AndroidEDU* program at 2009 providing an interesting interplay between the university and private IT firms as Google.

This course exhibited three stages: rise, stability and fall. The experience spanned six semesters, coinciding with the introduction, expansion and popularization of Android as a cutting-edge technology. The lessons learned along this period are summarized in a previous section.

The last stage of this course (accompanied with a loss of students' interest) has been attributed to a life-cycle that coincides with the standardization of cutting edge technologies, in this case of the Android system. Although this is a possible conclusion supported by the development of the course, other alternatives, such as lecturers' choices and actions taken to change and adapt the course to the new environment, may also explain the phenomenon. In such a case several open questions arise: how should lecturers and course designers adapt to the transformation of a cutting edge technology into a mainstream one? How should universities accompany this process? What role should IT firms play?

Although not conclusive, student assessments support the main features of the life-cycle. In the future, more accurate student assessments should be designed and applied for the duration of the course in order to be conclusive.

It is also claimed, that the life-cycle described in this work is common to every cutting edge technology course, but, a comparison with other cutting edge technologies courses is still to be done. This effort may be supported by knowledge on diffusion of innovations (Rogers, 2003).

One of the assets of the experience here described is the integrative nature of the contents of this course (also noted by (Muppala, 2011)) it remains to be explored how similar integrative benefits can be brought into main stream courses and with types of content.

REFERENCES

Ableson, F., et al. (2009). *Unlocking Android: A Developer's Guide*. Manning Publications.

Android Developers' Guide (2013). developer.android.com/guide.

AndroidEDU EMEA (2013). sites.google.com/site/androideduemea.

Bryan (2013). The rise and fall and rise of MOOCs. blogs.cofc.edu/tlt/2013/01/28/the-rise-and-fall-and-rise-of-moocs.

Burnette, E. (2009). *Hello, Android. Introducing Google's Mobile Development Platform*. The Pragmatic Bookshelf.

Frank, R. and Waren, A. (1982). Teaching new technologies. In *13th SIGCSE symposium on Computer science education*, page 268.

Google Play Store (2013). play.google.com/store.

Haseman, C. (2008). *Android Essentials*. Apress Ed.

HTC G1 (2013). www.htc.com/www/product/g1.

Ledford, J. (2009). *Web Geek's Guide to the Android-Enabled Phone*. QUE Ed.

Mednieks, Z., et al. (2012). *Programming Android: Java Programming for the New Generation of Mobile Devices*. O'Reilly Media.

Miller, M. (2008). *Google-pedia: The ultimate Google resource*. QUE Ed.

Muppala, J. K. (2011). Teaching embedded software concepts using Android. In *Proceedings of the 6th Workshop on Embedded Systems Education*, pages 32-37.

Murphy, M. and Allen, G. (2011). *Beginning Android 4*. Apress Ed.

Nelson, R. R. and Wright, G. (1992). The rise and fall of american technological leadership: The postwar era in historical perspective. *Journal of Economic Literature*, 30(4):1931–1964.

Nexus One (2013). www.google.com/phone/detail/nexus-one.

Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books.

Norman, D. A. (1998). *The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Solution*. MIT Press.

Postman, N. (1984). *Amusing Ourselves to Death. Public Discourse in the Age of Show Busines*. Penguin.

Rogers, E. M. (2003). *Diffusion of Innovations, 5th Edition*. Free Press.

Rogers, R., et al. (2010). *Android Application Development*. O'Reilly Media.

Developed and Taught Course Modules to Enhance Cloud Computing Education

Huiming Yu, Nakia Powell, Dexter Stembridge and Jonathan Rann
Department of Computer Science, North Carolina A&T State University, Greensboro, NC, U.S.A.

Keywords: Cloud Computing, Course Module, Hands-on Project, Education Tool.

Abstract: Cloud computing (CC) is emerging as a promising solution to businesses and individual users. In order to effectively teach knowledge of CC we have developed three course modules, one education tool, and one real world project. These modules have been successfully taught in COMP 621 Web Security course in Spring 2011, Spring 2012 and Spring 2013 in the Department of Computer Science at XXXX. Our experience exhibits that teaching these modules helped students not only understand the need of CC, but also gained significant knowledge of CC and security architecture. Through a laboratory exercise and a real world project students gained hands-on experience. Students' survey and feedback reflected that these modules and the laboratory exercise are very valuable in their educational experience. These modules and hands-on laboratory exercise could be used in software engineering, networking, Web security and other first year graduate level courses.

1 INTRODUCTION

Cloud computing provides flexible information technology infrastructures and configurable software services, reduces costs, enables business innovation and many others benefits. The service that cloud computing provides has become a commodity; when companies like Amazon, Intel, Google develop new applications, they sculpt cloud computing into the product they will produce. Several cloud computing systems have been developed by the industry, such as Amazon Elastic, Amazon EC2, Google Apps, etc. (Amazon Web Services, Google 2010). As a result expectations for Information Technology (IT) workforce have been increased by rapidly expanding cloud computing applications. It becomes an urgent task to teach knowledge of cloud computing to computer science students, so that they can meet the industry's expectation for the next generation IT professionals.

Most Computer Science, Information Management System, and Software Engineering curriculums do not adequately prepare students for such expectations. Based on demand three cloud computing course modules have been developed and taught, and the hands-on laboratory exercise has been designed and used in the Department of Computer Science at XXXX. In this paper, we present these new course modules titled "Introduction to Cloud Computing", "Introduction to Cloud Computing Services", "Security Architecture for Trustworthy Cloud Computing", one education tool, one hands-on laboratory exercise and one real world application project, and discuss our teaching experience. In section 2 the learning objectives and outcomes will be discussed. The details of these modules will be presented in section 3. A hands-on laboratory exercise will be specified in section 4. In section 5 a real world application project will be given. In section 6 teaching and evaluation results will be presented. The conclusion and future work will be discussed in section 7.

2 LEARNING OBJECTIVE AND OUTCOMES

Cloud computing is expanding rapidly. As a result it increases the expectations for IT professionals. In order to prepare students for IT workforce and for them to effectively learn knowledge of cloud computing we have developed three course modules. The learning objective of these modules is to enhance the computer science curriculum by providing information and technology that are related to cloud computing and cloud computing security. Instructors can teach these modules in

computer science senior classes or first year of graduate courses.

After attending the lecture of these modules, conducting a laboratory exercise and completing the assigned project students should be able to 1) understand the need for cloud computing, 2) gain significant knowledge of cloud computing architectures, characteristics and benefits, service models, deployment models, enabling technologies, data storage and operating system, and 3) apply learned knowledge to real world applications.

3 THE COURSE MODULES

Three courses modules have been developed that are "The Introduction to Cloud Computing", "Introduction to Cloud Computing Services", and "Security Architecture for Trustworthy Cloud Computing". We used three hours to teach the module 1, three hours for module 2 and two hours for module 3.

3.1 Module 1: Introduction to Cloud Computing

In this module, basic concepts, characteristic, benefits, enable technologies, service models and several cloud computing architectures are introduced.

In the basic concepts section we first answer the question why we need cloud computing. Cloud computing has the promise to provide all of businesses, or single user needs such as storage, computations, and applications in the mass, and the end user decides what they would like to have. With the increased demand for defining and molding cloud computing into something that is not that easily defined causes companies like Google, Salesforce, IBM to become very aggressive in terms of producing the product fast in order to be the inimitable leader. This causes a need for other companies like Twitter, Facebook to evolve and adapt in order to keep up with and hopefully exceed the pack of competing companies.

Secondly we briefly give the definition of clouding computing by introducing several definitions. One concept from Simon Wardley's "Cloud computing-Why it matters" presented at OSCON 09 was "Cloud Computing is a generic term used to describe the disruptive transformation in IT towards a service based economy driven by a set of economic, cultural and technological conditions" (Wardley, 2009). We use a graphic image to display the concept of cloud computing. Cloud computing can be defined as a paradigm in which a system utilizes the virtualized resources of networks of computers to handle a systems workload. A hypervisor must be a part of the system (Choi, 2009).

We present several existing systems to help students understand the architecture of cloud computing. First we introduce the architecture in a basic and simple way. A cloud computing environment consists of clients, the datacenter and the distributed servers (Velte et al., 2009). We use Amazon Elastic Computer Cloud (EC2) as an example because Amazon EC2 is a web service that provides resizable computing capacity in the cloud. It allows the user to obtain and configure capacity with minimal friction. It provides a user with complete control of their computing resources and lets them run on Amazon's proven computing environment. Amazon EC2 changes the economics of computing by allowing the user to pay only for capacity that they actually use (Amazon Web Services).

We discuss five main characteristics of cloud computing in this module. These characteristics are On-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. On-demand self-service implies a consumer can unilaterally provision computing capabilities. Broad network access has capabilities that are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms. Resource pooling gives the provider's computing resources which are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. Rapid elasticity can be quickly and elastically provisioned, in some cases automatically. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time. Measured service demonstrates cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service. Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service (Mell et al., 2009).

We also discuss three service models that are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) in this

module (Choi, 2009). SaaS gives users the capability to use the cloud provider applications running on a cloud infrastructure. PaaS gives users everything that they need to develop, test, and deploy applications to the cloud. IaaS is delivery of the computing infrastructure as a fully outsourced service, versus an in-house, capital investment-sourced model (Mell et al., 2009).

We introduce three key enabling technologies that are virtualization, hypervisor and Web services in this module. Over the last several years, virtualized technologies have become a standard in the world of computers. These technologies, including the virtual machine, are software implementations of computer systems which partition hardware into high-level abstractions. This allows for resource sharing within the cloud along with more affordable, flexible and scalable hardware services (Wang et al., 2008).

Virtualization can be looked at as the abstraction of computer resources from the actual computer which can successfully allow databases, middleware, and operating systems to practically run anywhere. Cloud computing paradigm allows workloads to be deployed and scaled-out quickly through the rapid provisioning of virtual machines or physical machines. Service virtualization combines and slices business services deployed independent of the operating systems, programming language or hosting location. An intermediary cloud gateway sits between the producer and the consumer and aggregates the Web Service Description Languages.

When software is run on a virtual machine, the bits representing the programs instructions run through a layer of software that simulates a dedicated server infrastructure. This layer of software is known as the hypervisor. Hypervisors allow servers to run a single instance of software that serves multiple clients at a time, without needing to rewrite the software. They also allow operating systems and applications to install a consistent hardware profile (Yu et al., 2012). Web services today are frequently just Application Programming Interfaces (API) or web APIs that can be accessed over a network, and executed on a remote system hosting the requested services.

3.2 Module 2: Introduction to Cloud Computing Services

In this module we discuss data storage as a service, software as a service, and cloud computing operating system.

Commonly known as Storage as a Service (StaaS) cloud computing presents a solution for any user's need with its robust infrastructure that far exceeds the needs of any user with an internet connection. StaaS allows users to store data at remote disks and access them anytime from any place (Wang et al., 2008). User's data is stored on remote storage systems that are maintained by a third party. The data is then accessible through the Internet.

Cloud storage as of today comes in many different architectural designs. A cloud based storage system connects the user to their stored information through a web application that allows the user to create, store, and file all the information chosen (Wang et al., 2008). This design shows where the user works on the application layer uploading files to the cloud, the software service then distributes the uploaded information to three separate storage locations for disaster recovery. The application, which is the web portal or interface to the cloud storage, allows users to subscribe to the web cloud storage. Copies of the files that the user submits are then transferred over the Internet to cloud storage servers.

We use Amazon S3 (Simple Storage Service) as an example to explain how StaaS works. Amazon S3 is a platform that allows for cloud storage to users on any scale from one gigabyte to thousands of exabytes. With low cost, scalability and reliability Amazon S3 can meet any software developers' needs [Amazon Simple Storage Service].

A cloud computing operating system (OS) provides an interface for developing and deploying massively scalable distributed applications on behalf of a large number of users and exploiting the seemingly infinite CPU, storage, and bandwidth. Traditional Operating Systems are used to manage hardware components within the computer. The cloud OS relies on distributed processes that manage the cloud resources. A cloud OS is similar as a traditional OS that is a collection of scheduler, virtual memory allocator, file system code, interrupt handlers, etc. to regulate the access by software to CPU, memory, disk, and other hardware peripherals. A cloud OS provides an additional set of functionalities to give administrative access to resources in the cloud. A cloud OS also provides a set of network-based interfaces that applications can use them to query the management system and control cloud resources (Pianese et al. 2010).

The Architecture of the cloud operating system consists of two separate spaces that are the User Space and Kernel Space (Pianese et al. 2010). The cloud's kernel space is used to regulate physical

allocation, access control, accounting and measurement of resources within the cloud's space. The cloud's user's space contains processes that are directly used by users.

Virtual Machine (VM) Management is the process that controls the allocation and de-allocation of virtual machines. Through the authentication kernel access rights are granted to the VM Management Kernal. This kernal is responsible for mapping an object name with its network address and port number. Catching all signals from the cloud OS, objects in the cloud user space are called and handled through this process.

Distributing processes is a key to provide abstraction and flexibility to computing resources from its physical nodes; this can be done with the use of Virtual Machines. Support for multiple OS's on the same hardware presents isolation to a certain extent between processes running on different VMs. Dynamic allocation on the cloud is done at two locations, the cloud process level migrates single processes between nodes while the VM level uses check pointing which can restore a VM state on a different node. Access control is controlled by a factotum server running on every machine, dedicated to authenticate the users resulting in a single sign on facility. The cloud user space contains libraries of API's that can be manipulated and changed to scale the application to satisfy different specifications just by updating a minor amount of embedded cloud library parameters (Pianese et al., 2010).

The cloud user space helps developers in two ways. One is eliminating the need for a developer to implement his or her own custom libraries, but instead use the cloud library API's and customize them. Secondly updating cloud library parameters for different specifications results in reusability of application components.

Software services known as SaaS (Software as a Service) identifies the use of computing power on a huge scale such as the cloud where companies can access applications and large amounts of virtual computing software without buying it. A service request is sent to the actual cloud service provider through a high level design. The web services architecture has three roles: a cloud provider, an application execution requestor, and a broker/mapper. The provider creates the web service and makes it available to clients who want to use it. An application requestor is a client application that consumes the web service. The mapper provides a way for the provider and the requestor of a web service to interact. The provider, requestor, and broker interact with each other through the operations of publish, find, and bind. A provider informs the broker about the existence of the web service by using the broker's publish interface to make the service accessible to clients. The information published describes the service and specifies where the service is located. The requestor consults the broker to locate a published web service (Lu, 2010).

3.3 Module 3: Security Architecture for Trustworthy Cloud Computing

Many security technologies have been proposed and developed to enhance cloud computing security. In this module we use Cisco Secure Data Center Framework as an example to discuss security issues. Cisco Secure Data Center Framework is one of them that provides multiple security layers and applied different existing security technologies to enhance cloud computing security [Bakshi].

Cisco has developed the Cisco Secure Data Center Framework with many security and trust considerations. The first consideration is general security issues. In this framework traditional security issues of information assurance such as data access control, encryption, and incident detection are integrated. The second consideration is access control that means an enterprise has capability to directly manage how and where data and applications are deployed and used. The third consideration is compliance and service-level management (SLA). It includes contracting and enforcement of service level agreements between different parties, legal issues, regulation and industry requirements, etc. The framework consists of Threat Profile, Cloud Data Center Visibility, Cloud Data Center Protection, Cloud Data Center Building Blocks, Cloud Data Center Control, Cloud Data Center Compliance and SLA. The Threat Profile contains threat models such as Service Disruption, Intrusion and Takeover, Data Leakage, Data Disclosure, Data Modification, Identity Theft and Fraud, etc. The Cloud Data Center Visibility provides various functions such as intrusion detection, anomaly detection, packet capture, network data collection and monitoring as well as event analysis and correlation. The Cloud Data Center Protection is very straightforward and provides different protection techniques and mechanisms such as stateful firewall access control, intrusion prevention, content filtering, and enforces endpoint and baseline security. Cloud Data Center Control deals many issues with how data being

control. Because data are centralized in Cloud Computing it increase the possibility of insider threats, therefore a compartmentalization strategy is important. The data encryption policy must be in place for customers and providers. For an administrator, who accesses and controls virtualized operating system, must be strong authenticated. The Cloud Data Center Compliance and SLA deal with how the data being handled is controlled. So certain requirements are set in place by the consumer and provider in order to reach an agreement to make sure requirements for both are met. The service level agreement is in place to enforce these requirements and to make sure that the consumer and provider remain compliant [Bakshi].

4 A HANDS-ON LABORATORY EXERCISE

A key issue for students learning cloud computing knowledge and apply it to real world applications is to get hands-on experience. We have developed a deployment education tool and a one hour laboratory hands-on exercise to help students learn related knowledge and use cloud computing platforms.

4.1 Using Microsoft Windows Azure Platform

There exist many cloud providers. All of which have platforms associated with the deployment of applications to their cloud environment. Each provider implements its cloud service in order to cater to the needs of clients.

We selected Microsoft Windows Azure as our cloud platform. This platform provides and maintains control of an integrated development and hosting environment to allow software developers to easily create, manage, scale, and host applications through Microsoft datacenters. Applications are executed using Azures compute services, while storage is handled by the storage services; both of which are built on top of a component known as the Windows Azure fabric, which ties all services together and monitors applications to maintain desired states (Chappelle, 2008). These services are the main components of Windows Azure, shown in Figure 1. Specifically, a few of the responsibilities of the Azure fabric is to create new VM's and assign them to processor cores, run instances of an application, and detect when a instance has failed.

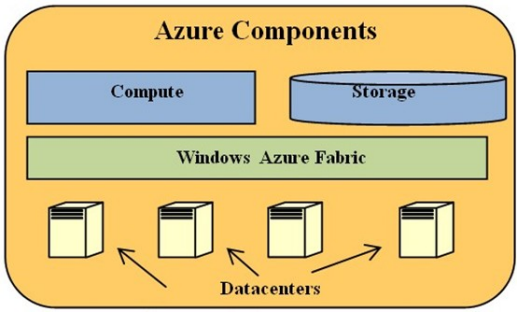


Figure 1: Windows Azure components.

4.2 Designing and Implementing an Education Tool

We designed and implemented an education tool named Scientific Utility for Cloud Harnessing (SUCH) (Yu et al. 2012). This tool provides functions to successfully deploy scientific applications to the Windows Azure cloud platform in a less complex manner than other methods, thus helping students to get hands-on experience of using cloud services. SUCH consists of three main components that are the deployment manager utility, the message queue, and the worker role. These components are shown in Figure 2. All of these components work together to upload, execute, and manage deployments within Windows Azure.

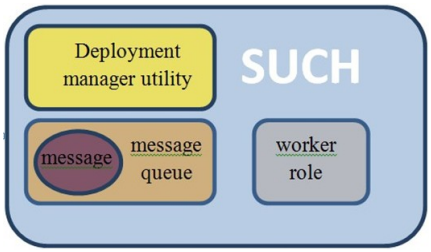


Figure 2: SUCH components.

The deployment manager utility allows students to upload files as well as manage services and applications utilizing the cloud. When the deployment manager is executed, it initiates cloud components, uploads files, and creates, starts, stops, and deletes deployments. To execute it, users must navigate to the Azure Application executable file located in Azure Deployment Manager folder in SUCH. The Message Queue utility builds messages, and stores them on the message queue for the worker role. It is composed of three programs: RunScienceCodeMessage.cs, a bucket class used to create message objects for the message queue, PopulateMessage.cs, which creates message object

instances, initiates the message queue in Azure, serializes messages and populates the message queue, and ClearMessageQueue, which clears the message queue of messages. The Worker Role, located in the Application Worker Role folder, first uses an infinite loop to check the queue for instructions. If a message is located, the worker role takes it from the queue and executes message instructions with the assistance of Azure components.

SUCH provides a simple means of utilizing the vast resources within the cloud, thus students can avoid many unnecessary complications. It is user-friendly; students can join the laboratory lecture, read a few pages of instructions, complete set up and configuration steps, then they can execute their applications in the cloud.

4.3 Developing a Hans-on Laboratory Exercise

We developed a one hour hands-on laboratory exercise to help students understand how Microsoft Windows Azure works, how to develop and deploy an application to a cloud system. We wrote a document titled "User Manual for Using Microsoft Windows Azure and SUCH". It describes all details of ten steps to guide students using Microsoft Windows Azure and SUCH.

First students must sign up for an account and downloading the Windows Azure SDK. Then students should download the SUCH zip file containing Clear Message Queue, Populate Message Queue, Application Worker Role, and Azure Deployment Manager. Students need to create an IIS management certificate and upload it to Azure via the management portal. They must fill in storage credentials into the values in the setting property of the SUCH worker role. Students need to generate a package file via the publish function in the SUCH worker role. Once the file is generated, they navigate to the Publish folder in the Application Worker Role folder, copy the configuration and service package files, and move them into the Debug folder of the Deployment Manager Utility Folder. Once those steps are completed, students must fill in account name, account key, etc. for the configuration files in the populate message queue, clear message queue, and deployment manager utilities. The file intended to be executed in the Azure should then be uploaded to Azure storage. Finally, users simply run the AzureDeploymentManager.exe, which will progress through a sequence of steps which enable successful application deployment. As SUCH executes, students can navigate to the online Azure management portal to view the creation, starting, stopping, and deletion of deployments. Figure 3 shows the Azure management portal as a deployment is in its Starting phase.

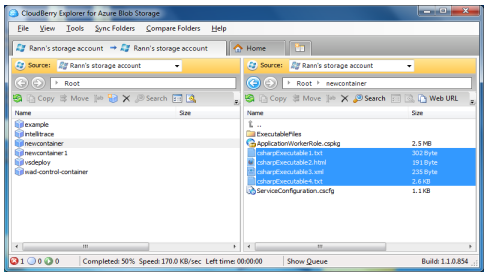


Figure 3: Output files created from executed applications within Azure.

In order to deploy to Azure, all executable applications must be uploaded to blob storage using CloudBerry Explorer for Azure Blob Storage. Using CloudBerry Explorer, a student would navigate to the Release folder location in the bin section of each folder that the executable file is located in. Next the student would execute SUCH and wait until the deployment has fully started. Once the deployment is ready, the student would fill the Populate Message Queue configuration file with information specifying the name of the executable files to be executed in Azure, location of each of the executable files in Azure storage, and the expected output file. After SUCH successfully runs, studnets should check their storage container to see if a new output file was created and uploaded to the cloud storage.

4.4 Getting Hands-on Experience

After students attend the lectures of the new course modules they are required to attend the one hour hands-on laboratory exercise. In the laboratory, we briefly introduce the functions of Microsoft Windows Azure and how to use the Scientific Utility for Cloud Harnessing. Then we guide students to go through installation and create a configuration file. We have developed several small application examples to let students to deploy them and to run these examples in Microsoft Windows Azure according to the steps that are described in 4.3. One application example is that prints a message to a text output file and fills an array with random numbers, then sorts the array in ascending and descending order, and displays the results in the output file. The results are shown in Figure 4. The output text file

first displays a message explaining the functions of the program such as 'creates an array of a user specified number of elements'. The next line displays an array of 50 elements, containing random numbers between 0 and 1000. The line after that displays the sorted array in ascending order, and the last line displays the sorted array in ascending order.

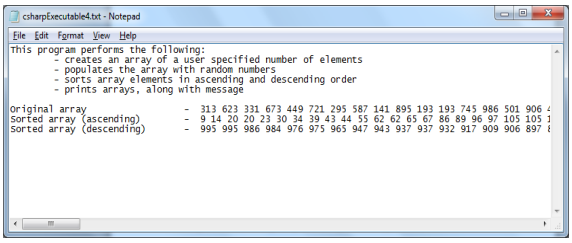


Figure 4: The output file.

Because COMP 621 Web Security is a graduate class most students (about 80%) can follow the instruction to deploy and run the sample program successfully. Some students need Teaching Assistants help.

5 A REAL WORLD APPLICATION PROJECT

Another key issue for students learning cloud computing knowledge and applying it to real world is to let them develop an application and run it in an existing cloud computing platform. Using this method students learn how to solve a real world problem and how to use a cloud computing system, and reduce the gap between classroom and the real world.

We assigned a real world application project to students. Two or three student formed a team. We provided detailed information about Microsoft Windows Azure, such as Azure account management, Azure services, configuration and package files, etc. to students. The project consisted of four major phases. In phase I each team developed a real world application program and executed it on a local computer. The program could be a hotel reservation system, a car dealer system, a student information management system, etc. In phase II students used SUCH to deploy the developed application to the Microsoft Windows Azure. In phase III students executed the application in the Microsoft Windows Azure and putted output information into a data file. In phase V students checked their storage container (via Cloudberry) to see if a new data file was created and uploaded to storage in the cloud. Each team must hand in a project report that included how the program was developed, how they deployed the program to the cloud, and the executed result. We also required every team to demonstrate their project in the class and exchange experience with other students within ten minutes.

6 RESULTS

We have successfully taught three course modules in COMP 621 Web Security class in the Department of Computer Science at XXXX in Spring 2011, Spring 2012, Spring 2013 and received excellent results.

Eighteen students took this course in Spring 2011, twenty-six in Spring 2012 and sixteen in Spring 2013. The students were very excited about learning the new emerging technology of cloud computing. To evaluate the students' reactions and get feedback for teaching these modules we conducted a pre and post surveys at each semester. These surveys were in addition to the university course evaluation. The survey consists of two groups. The first group contains eight questions that reflected knowledge level of students. These questions are 1) basic knowledge of cloud computing, 2) knowledge of characteristics, service models and enable technologies, 3) knowledge of architectures, 4) knowledge of data storage services, 5) knowledge of operating systems, 6) knowledge of application software services, 7) knowledge of threats, and 8) knowledge of security mechanisms. Students' ratings could be very low, low, medium, high, and excellent. The Spring 2013 survey results are displayed in Figure 5. Blue color represents pre-survey and red color represents post survey. For the pre-survey majority of the students gave a very low rating for their knowledge of characteristics, service models, enable technologies, architectures, deployment models, application software services and security mechanisms of cloud computing. The results varied more for students' knowledge on data storage services, operating systems and threats of cloud computing, while majority of students seemed to have basic knowledge of cloud computing. No students rated their knowledge as High for any questions in the pre-survey. The post survey results show that students learned cloud computing and improved their knowledge levels. They had big improvement in their knowledge of characteristics, service models, enable technologies, architectures, data storage services, operating systems, threats, and security mechanisms of cloud computing. The

second group are general questions that are 9) these modules help you gain knowledge of cloud computing and security related issues, and 10) you are more knowledge about cloud computing, threats and security mechanisms. Students could select Strongly Agree, Agree, Neither Agree or Disagree, Disagree, or Strongly Disagree. 100% students selected agree or strongly agree for all questions in the second group. Students felt these modules stimulated their interest to learn more about cloud computing. Most students enjoyed the hands-on laboratory exercise.

7 CONCLUSIONS

This paper presents three new cloud computing course modules titled "Introduction to Cloud Computing", "Introduction to Cloud Computing Services" and "Security Architecture for Trustworthy Cloud Computing", an education tool, a hands-on laboratory exercise and a real world application project assignment. These modules cover the basic concepts, architecture, characteristics and benefits, service models of cloud computing. These modules also discuss virtualization, hypervisor, Web services of enabling technologies, deployment models and cloud computing operating systems. Microsoft Windows Azure was selected as the cloud platform for the hands-on laboratory exercise. An education tool, Utility for Cloud Harnessing was developed to provide functions to deploy applications to the Windows Azure cloud platform in a simple manner, was used. A hands-on laboratory exercise has been developed to guide students to deploy and run an application in the Azure cloud platform step by step. Finally a real world application project was assigned to students.

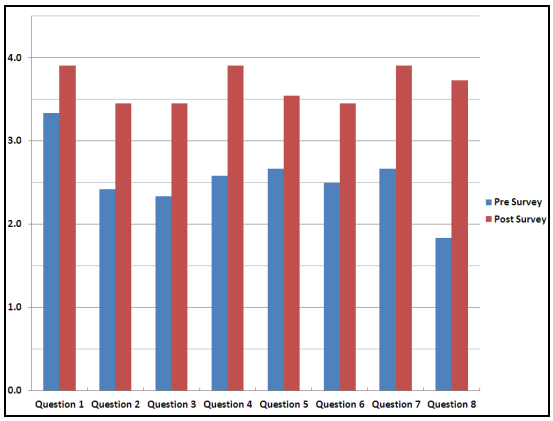


Figure 5: Spring 2013 Surveys.

We have successfully taught these modules in COMP 621 Web Security course in Spring 2011, Spring 2012 and spring 2013. Students' survey and feedback reflected these modules are very valuable. They felt the module stimulated their interest to learn more about cloud computing. Our experience exhibits that combining lectures and hands-on laboratory exercise together can successfully integrate education and real world application into cloud computing education. The lectures bring new technology to students; let them learn the knowledge and applications of cloud computing. The hands-on laboratory exercise demonstrates how to deploy and run an application in an existing cloud platform, helps students understand functions of a real world cloud computing system and learn how deploying and running an application on the cloud platform, and lets students gain important insights into how applying learned knowledge into a real world. Assigning a real world application as a project helps students understand the difference running an application on a local computer and on a cloud computing system, increases their capability to solve complicated problems and breaks barriers between education and real world applications. Our experience also proves that teaching cloud computing modules is beneficial to students through both theory and hands-on experience, helps them learn advanced technology and apply learned knowledge to real world applications.

In order to satisfy the high demand of cloud computing professionals we are developing a course titled "Cloud Computing and Security". We will offer this course in Fall 2014.

ACKNOWLEDGEMENTS

This work was partially supported by National Science Foundation under the award number CNS-0909980, DUE-0830686 and DUE-1129136.

REFERENCES

Amazon Web Services, DOI=http://aws.amazon.com/ec2/

Google 2010 Google Apps, DOI=http://code.google.com/apis.

Wardley, S., 2009. Cloud Computing-Why it Matters. *O'Reilly Open Source Convention*.

Choi, P., 2009. How to Develop an Effective Security Strategy to Play in the Public Cloud. *Cloud Computing Journal.*

Velte, T., Velte, A. and Elsenpeter, R., 2009. Cloud

Computing A Practical Approach. McGraw-Hill.

Mell, P. and Grance, T., 2009. The NIST Definition of Cloud Computing. Version 15, October 7, 2009.

Chappelle, D., 2008. Introducing the Azure Services Platform. *Microsoft Corporation.*

Wang, L., Tao, J. and Kunze, M., 2008. Scientific Cloud Computing: Early Definition and Experience. *In Proceedings of 10th IEEE International Conference on High Performance Computing and Communications.*

Wu, J., Ping, L., Ge, X., Wang, Y. and Fu, J., 2010. Cloud Storage as the Infrastructure of Cloud Computing. *In Proceedings of IEEE Int. Conf. on Intelligent Computing and Cognitive Informatics*.

Amazon, Amazon Simple Storage Service, DOI=http://aws.amazon.com/s3/faqs/#What_can_I_do_with_Amazon_S3.

Pianese, F., Bosch, P., Duminuco, A., Janssens, N. Stathopoulos, T. and Steiner, M., 2010. Toward a Cloud Operating System. *In Proceedings of IEEE Network Operations and Management Symposium*.

Lu, X., 2010. An Approach to Service and Cloud Computing Oriented Web GIS Application. *In Proceedings of IEEE International Conference on Internet Technology and Applications*.

Bakshi, K. Cisco Cloud Computing - Data Center Strategy, Architecture, and Solutions. DOI=http://www.cisco.com/web/strategy/docs/gov/CiscoCloudComputing_WP.pdf .

Yu, H., Rann, J. and Zhan, J., 2012. SUCH: A Cloud Computing Management Tool. *In Proceedings of the 5th IFIP International Conference on New Technologies, Mobility & Security*.

Better Feedback for Educational Online Judges*

Anaga Mani[1], Divya Venkataramani[1], Jordi Petit[2] and Salvador Roura[2]
[1]*SASTRA University, Tanjore, India*
[2]*Computer Science Department, Universitat Politècnica de Catalunya, Catalunya, Spain*

Keywords: Online Programming Judges, Automatic Assessment, Data Mining.

Abstract: The verdicts of most online programming judges are, essentially, binary: the submitted codes are either "good enough" or not. Whilst this policy is appropriate for competitive or recruitment platforms, it can hinder the adoption of online judges on educative settings, where it could be adequate to provide better feedback to a student (or instructor) that has submitted a wrong code. An obvious option would be to just show him or her an instance where the code fails. However, that particular instance could be not very significant, and so could induce unreflectively patching the code. The approach considered in this paper is to data mine all the past incorrect submissions by all the users of the judge, so to extract a small subset of private test cases that may be relevant to most future users. Our solution is based on parsing the test files, building a bipartite graph, and solving a Set Cover problem by means of Integer Linear Programming. We have tested our solution with a hundred problems in Jutge.org. Those experiments suggest that our approach is general, efficient, and provides high quality results.

1 INTRODUCTION

It is an established fact that practice is fundamental to learn computer programming. Whether at secondary, high school or university level, instructors usually assign programming problems to beginner students in order to help them acquire this skill. However, it is unquestionable that, for instructors, correcting programming assignments can be a slow, boring and error-prone task. Fortunately, programming problems are ideal candidates for automated assessment; see (Douce et al., 2005) for a review on this topic, (Ihantola et al., 2010) for an updated list of such systems and (Kurnia et al., 2001; Cheang et al., 2003; Joy et al., 2005) for some experiences in the classroom.

Currently, programming online judges are widely used to assess solutions to programing problems. Online judges are web based systems that offer a repository of programming problems and a way to submit solutions to these problems in order to obtain an automatic verdict on their behavior when run on different public and private test cases designed to be as exhaustive as possible. The solutions of such problems are complete programs or functions with a well defined specification.

Historically, online judges were targeted to train participants for important programming contests as the IOI[1] or the ACM-ICPC[2]. UVa[3], Timus[4], Sphere[5] and Codeforces[6] are just a few examples of such judges. More recently, some judges as TopCoder[7] and InterviewStreet[8] were built as a recruitment platform where companies can find and try to hire highly skilled programmers.

Additionally, some online judges were designed with an educative purpose in mind. Some notable examples are EduJudge[9], URI Online Judge[10] and Jutge.org[11]. These systems offer an organized repository of graded problems, and integrate the features of an online judge with those of a learning management system to assist both students and instructors. We refer to (Verdú et al., 2012; Tonin et al., 2012; Petit et al., 2012) for an overview of their respective capabilities, and to (Giménez et al., 2009) for an experience on the usage of Jutge.org in a first-year in-

*Work supported by the Spanish project FORMALISM (TIN2007-66523) and by the Generalitat de Catalunya's project ALBCOM (ref. 2009 SGR 1137).

[1]http://www.ioinformatics.org.
[2]http://icpc.baylor.edu.
[3]http://uva.onlinejudge.org.
[4]http://acm.timus.ru.
[5]http://www.spoj.pl.
[6]http://codeforces.com.
[7]http://www.topcoder.com.
[8]http://www.interviewstreet.com.
[9]http://eduvalab.uva.es/en/projects /edujudge-project.
[10]http://www.urionlinejudge.com.br.
[11]http://www.jutge.orgJutge.org.

troductory programming course.

A common feature of online judges is that their verdict is, essentially, binary: Either the submitted solution is correct or it is not. Clearly, these binary verdicts hinder the adoption of online judges in educative settings. Indeed, some students may feel frustrated when their solution is evaluated as wrong but no example where it fails is offered by the system. Likewise, some instructors may be reluctant to use a tool for which they cannot access the private test cases. Since beginners tend to write overly complicated solutions, online judges often detect errors that can be very hard to find by just inspecting the source code, even for experienced instructors.

As long as online judges require to keep secret their private test cases, there does not seem to be much to do to increase the feedback. However, in some particular settings it is feasible to automatically correct some wrong solutions (Singh et al., 2013). Another possible approach, the one used in this paper, is to relax a bit the secrecy and consider that, for educative purposes, it could be positive to reveal a subset of the private test cases.

An obvious solution would be to just show users an instance for which their code fails. However, we feel that just providing an arbitrary example where a program fails would encourage sloppy-thinking programmers, who first recklessly write a wrong program, and afterwards try to patch it (several times) using the given counterexamples, which, worse still, could be too insignificant again. Clearly, this is not the right way to proceed. Instead, we believe that the counterexamples provided (at least to the instructors, and perhaps also to the students; see the conclusions of this paper) should be as relevant as possible. By "relevant" we mean that these test cases should probably be helpful to most of the users. The goal of this paper is to show how to create these relevant test cases.

Our solution, which we call *the distiller*, consists in data mining past incorrect submissions sent by all the users of an online judge. We will present some examples that indicate that this "distillation" process, which we have implemented in Jutge.org, is quite general and efficient. Although we lack a quantitative measure for the quality of the produced test cases, we will show that, in general, a set of a few small test cases is enough to make (almost) *all* bad submissions fail. Therefore, it could be argued that those *automatically generated* test cases capture somehow the essence of the problem, and thus should be of great help when revealed to some users.

To the best of our knowledge, this is the first time that such a study has been conducted.

Organization. The paper is organized as follows: In Section 2, we give a more detailed overview of the main elements of online judges and present some basic definitions. The basic description of the problem we tackle is given in Section 3, together with the main ingredients of our solution. An overview of the results obtained by our tool for a few selected problems from Jutge.org is shown in Section 4. Section 5 discusses how to use the distilled test cases.

2 BACKGROUND

In this section, we first summarize the main elements and resources of online judges, then properly define the different types of test cases and submissions we are interested in.

Main Elements of Online Judges. Although online judges have differences among them, they all have a common set of elements and resources and a common way to interact with them. On one hand, judges keep information on their users and provide functionality to register, log in, etc.

On the other hand, judges offer a repository of problems. Problems consist of a statement (typically in HTML or PDF format), some test cases, and some internal specification on how to run the test cases under certain constraints as time or memory limits.

The basic interaction between a user and the judge is the user submitting a solution to some problem in the repository. At this point, the judge will queue the submission to check for its correctness using the test cases and, subsequently, the judge will provide a verdict to the user.

Processing a submission is a somehow difficult task, because it must be performed in a secure way that defends the judge against possible attacks. We shall ignore these important implementation topics as they are not relevant to the current work, but refer the interested reader to (Leal and Silva, 2003; Forišek, 2006; Petit et al., 2012).

Classification of Test Cases. Besides the statement, problem setters define two kinds of test cases:

- *Public Test Cases*. Test cases that are visible by the users. They are usually included in the statement of the problem, in order to clarify it and to provide a few examples of the expected input and output.
- *Private Test Cases*. Internal test cases that are never disclosed to the users. They are used to check, in the most possible exhaustive way, the

correctness of the submissions. These tend to include hand made tricky inputs and automatically generated inputs.

Most problems expect solutions that read from the standard input and write to the standard output. Thus, the different test cases are stored in pairs of files, one for the input and another for the correct output.

Because of practical reasons (for instance, the overhead of creating a new process to do very small work), many individual test cases are usually joined in a few files. To clarify this, consider the problem of computing the greatest common divisor of two numbers. Assume that this problem has n individual test cases (say, $n = 10^5$). Then, instead of having n input files and their corresponding n output files, we rather have a single input file with the n pairs of numbers and one output file with the n correct answers. Of course, the problem statement asks to read a sequence of pairs of numbers and, for each one, write their gcd. Moreover, in addition to this big file, the problem setter will often add a few small files, to check, e.g., the correct processing of a file with just one pair of numbers or with no numbers at all.

In the following, we will stress this distinction by denoting by *test cases* the individual tests that are joined in a *test file*. In general, problems have a few test files (say, at most a dozen) with many test cases (say, some thousands). When needed, we will also distinguish between *input test files* and *correct output test files* and between *input test cases* and *correct output test cases*. Likewise, we will use the terms *user output test files* and *user output test cases* for the outputs generated by a (possibly incorrect) submission from a user.

Classification of Submissions. Online judges traditionally classify submissions with a marking system that is mainly binary. A solution is either *"Accepted"* (AC) when it efficiently produces satisfactory results on all the test cases run by the judge, or it is *"Rejected"*. Despite of that, *"Rejected"* submissions usually carry some additional information to provide a better verdict: *"Compilation Error"* (CE)[12], meaning that the submitted code cannot even be successfully compiled, *"Execution Error"* (EE) meaning that some error appeared during the execution of some test case (possibly reporting a refined diagnostic as division by zero, invalid memory access, memory limit exceeded, or the most common time limit exceeded...) or *"Wrong Answer"* (WA), meaning that, despite a successful execution, the user output does not match the correct output. This last verdict indicates a solution that is fast enough and does not crash, but which is wrong. Note that the reason for a WA verdict can range from a subtle bug in the code to a program that has little or no sense at all. In any case, the contents of the private test cases are never disclosed to the user.

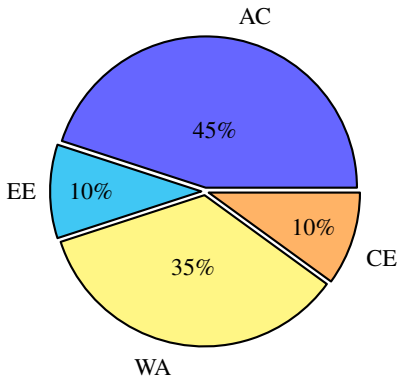


Figure 1: Current distribution of verdicts for Jutge.org.

Figure 1 shows the current distribution of verdicts for Jutge.org.

For the purpose of this work, we are interested in an alternative classification of submissions:

- *"Good Submissions"*. Submissions that pass all test cases, public and private. These correspond to *"Accepted"* submissions.
- *"Bad Submissions"*. Submissions that pass all public test cases but fail in some private test cases.
- *"Horrible Submissions"*. Submissions that do not even pass all public test cases.

In the rest of the paper we shall not be interested in good submissions (because, being good, they do not help us to find relevant test cases) nor in horrible submissions (because these are so bad that they do not reflect a minimum of quality). Thus, we focus in bad submissions.

Please observe that the *good, bad* and *horrible* terms are labels that we only associate to certain types of submissions and should not be applied to their authors. Moreover, these terms exclusively apply to the black-box correction process of the judge, which is based on the functional correctness and efficiency of the submitted codes, but not on other important qualities, such as readability, succinctness, elegance, etc.

[12] See, e.g., https://www.jutge.org/documentation/verdicts or http://uva.onlinejudge.org/index.php?option=com_content&task=view&id=16&Itemid=31.

3 PROBLEM DESCRIPTION AND SOLUTION

Consider any particular problem in the repository of an online judge. Our goal is to extract a relevant test file with a small set of private test cases for which all bad solutions fail for at least one test case in the test file. For simplicity, we will generally assume that the number of test cases in the relevant test file must be minimized. When necessary, we will point to the minor changes to apply so to reduce the total size of the file itself.

In order to discover these relevant data sets, we assume that the following information is available:

- A set of private input test files together with their corresponding correct output test files.
- A "large enough" collection of bad submissions.
- For each input test file and each bad submission, the corresponding user output file.

In the case of Jutge.org, all this information is permanently stored in the system databases, but in other judges the user output files are not kept and should be regenerated by resubmitting the bad submissions.

Let n denote the total number of test cases for the problem, and let m be the total number of available bad submissions.

Our solution is conceptually described in Figure 2 in the next column. In the following, we provide further details for some of those steps.

Splitting Files (Steps 1, 2 and 3). To split an input test file consists in parsing it in order to extract its individual input test cases. Output test files are split in a similar way.

The splitting phase is highly dependent on the specific syntax prescribed for the input and output files. In the worst situation, an ad-hoc program to parse the input and the output is necessary for each problem. Fortunately, a large fraction of elementary problems share the same types of input/output general parsing schemes, possibly involving a few parameters. Of course, each individual problem must still be tagged by hand to specify how to split its inputs and outputs.

These are the most frequent parsing schemes that we have identified among those implemented in our tool:

- *No split*. No split needs to be done, as all the input in each test file represents a unique test case.
- *Split by lines*. Each line of the file corresponds to an individual test case.

1. Split all input test files into a list of input test cases $[I_1, \ldots, I_n]$.
2. Split all correct test files into a list $[C_1, \ldots, C_n]$, so that C_i is the correct output for I_i.
3. For each bad submission $1 \leq j \leq m$, split all its user output test files into a list $[O_1^j, \ldots, O_n^j]$, so that O_i^j is the output of the j-th bad submission for the i-th test case.
4. Build a bipartite graph:
   - Its n left vertices correspond to each test case.
   - Its m right vertices correspond to each bad submission.
   - For each submission j that fails on a test case i, (i.e., $O_i^j \neq C_i$), add an edge between i and j.
5. Compute a minimum subset $S \subseteq \{1, \ldots, n\}$ of left vertices whose induced edges hit all right vertices.
6. Output all the I_k, for all $k \in S$, as the relevant test cases for the problem.

Figure 2: *Distiller* algorithm.

- *Split by words*. Each word of the file corresponds to an individual test case. Here, a word is understood as a sequence of non-blank characters.
- *Split by K words*. Each K consecutive words of the file corresponds to an individual test case.
- *Split with counter*. Each test case is made up of several elements (lines, words, K words,...) and the number of them is specified with an initial counter.
- *Split with separator*. Each test case is made up of several elements (lines, words, K words,...) and they are separated with a special separator to be specified as a parameter.

In the case that a problem needs a non predefined splitting scheme, our implementation allows us to define a customized scheme (see Section 4).

Solving the Set Cover Problem (Steps 4 and 5). Step 4 builds a bipartite graph where left vertices corresponds to test cases, right vertices to submissions, and edges correspond to failure to report the correct output for a particular submission on a particular test case. Step 5 asks for solving the following problem: Given a bipartite graph $G = (U, V, E)$, find a minimum subset S of U such that each vertex in V has some neighbor in S. Figure 3 illustrates it.

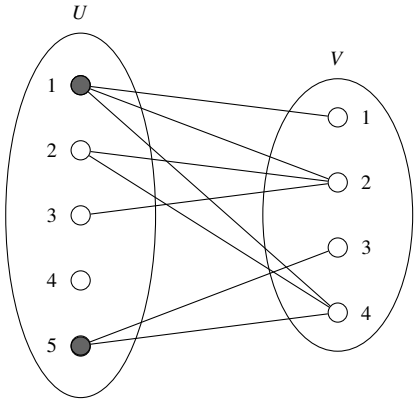


Figure 3: The Set Cover problem: Given a bipartite graph $G = (U, V, E)$, find a minimum subset S of U such that each vertex in V has some neighbor in S. In this example, the minimum solution is obtained with $S = \{1, 5\}$.

It turns out that this problem is one of the formulations of the *Set Cover problem*, which is a classical problem in computer science. Its decisional version is **NP**-complete, as proved in (Karp, 1972). Some heuristics have been considered to efficiently obtain solutions of good quality for this problem. It is known (Chvatal, 1979) that the obvious greedy algorithm (at each stage, choose the test case connected the largest number of still uncovered submissions) has a logarithmic approximation ratio. Unfortunately, this is essentially the best possible ratio under the usual $\mathbf{P} \neq \mathbf{NP}$ assumption (Alon et al., 2006).

Despite the theoretical intractability of Set Cover, we are able to exactly solve our instances using Integer Linear Programming (ILP). Let x_i be the binary variable that encodes whether $I_i \in S$ or not. Also, let $E_{i,j} \in \{0, 1\}$ indicate whether there is an edge connecting the i-th test case with the j-th submission. Then, the set cover problem is modeled as

$$\text{minimize} \sum_{i=1}^{n} x_i$$

subject to

$$\sum_{i=1}^{n} x_i \cdot E_{i,j} \geq 1 \qquad \forall j = 1, \ldots, m$$
$$x_i \in \{0, 1\} \qquad \forall i = 1, \ldots, n.$$

Output Distilled set (Step 6). In order to output the input test cases that make up the distilled set, these must be joined in the opposite way that they were split. To do so, each parsing scheme must implement this reverse operation.

Further Details. Let us finally point to some additional details relevant to our implementation:

- To speed-up the running time of our solution, our complete implementation does not work with the output test cases themselves, but with numeric hashes instead. This enhancement keeps true negatives, but can introduce false positives (a wrong test case might be deemed correct) with a tiny probability, which we decide to ignore for the sake of efficiency.
- In some rare cases, we have found that some bad submissions are "*really bad*". These are detected by observing that the number of user output test cases is different than the number of correct output test cases or, when being equal, these differ in too many positions. To avoid junk into the graph, our algorithm ignores these submissions in a parameterizable way.
- In order to optimize the total size of the distilled set, we can simply change the objective function of the ILP so to weight each input test case by its size. Other variations such as maximizing the number of bad submissions caught by a distilled set of a given maximal size could be defined in the same style.

4 RESULTS

We have implemented a program in Python using a commercial product (Gurobi, 2013) to solve the ILP.

For testing, we have chosen about a hundred different problems from the introductory programming course of Jutge.org and their corresponding submissions of about 7200 users since September 2006. In the sake of providing results on the most popular problems, the selection was based on their number of submissions.

Each of these problems with identifier *id* is freely available at http:/www.jutge.org/problems/*id*. All statistics are at the time of writing this paper.

Validating Dates (P29448). This problems asks for a program that reads a list of triples of integers and, for each one, prints whether it is or not a valid date in the Gregorian calendar. Its public test cases are shown in Table 1(a).

Any experienced programmer would agree that this problem is easy but presents a few complications for a beginner, because it requires a thorough case analysis for the validation on the number of days for each month, with the additional difficulty of copying with leap years.

This problem has five private test files, which feature a total of 129458 elemental test cases (128163

Table 1: Results for problem "Validating dates".

(a) Public samples

Sample input	Sample output
30 11 1971	Correct Date
6 4 1971	Correct Date
4 8 2001	Correct Date
29 2 2001	Incorrect Date
32 11 2005	Incorrect Date
30 11 2004	Correct Date
-20 15 2000	Incorrect Date

(b) Statistics

Submissions	6903
Accepted	1600
Rejected	5303
Compilation Error	378
Wrong Answer	4856
Execution Error	69
Users with accepted problem	1115
Users with rejected problem	276
Good submissions	1600
Bad submissions	3718
Horrible submissions	1585

(c) Distilled input test cases

29 2 4900	13 9 1800	29 2 1870
7 2 1900	28 0 1852	31 2 1964
31 8 8298	31 1 7709	1 12 2002
29 1 1942	31 12 5741	30 1 3157
30 2 1864	29 2 6592	0 2 1910
31 7 3535	31 9 1948	29 2 2000
31 2 2000	32 12 2008	0 3 2000
5 13 2000	31 11 2000	

without duplicates). The statistics for this problem are shown in Table 1(b).

The parsing scheme for this problem was set to *split by 3 words* and the parsing scheme for the output was set to *split by lines*. The application of the distillation process took about 35 seconds (including 1.10 seconds to exactly solve the Minimum Set Cover problem with ILP) and produced a distilled set with 23 test cases, shown in Table 1(c).

As could be expected, those test cases feature limiting values for the number of the day (such as 0 and 32) and stress years (such as 2000) for which the leap year rule may be wrongly coded.

Roman Numbers (P18298). This problems asks for a program that reads several (decimal) numbers and prints their equivalent Roman numbers. All numbers are in the range $[1,3999]$, and the rules to transcribe to Roman numbers are given. Its public test cases are shown in Table 2(a).

Table 2: Results for problem "Roman numbers".

(a) Public samples

Sample input	Sample output
1	1 = I
4	4 = IV
10	10 = X
16	16 = XVI
2708	2708 = MMDCCVIII
999	999 = CMXCIX
3005	3005 = MMMV

(b) Statistics

Submissions	2338
Accepted	1352
Rejected	986
Compilation Error	178
Wrong Answer	758
Execution Error	50
Users with accepted problem	1087
Users with rejected problem	39
Good submissions	1352
Bad submissions	867
Horrible submissions	119

(c) Distilled input test cases

2490	3470	1000	3049

This problem has three private test files, which feature 127977 elemental test cases (but just 3999 without duplicates; the remaining ones are there just to ensure that the programs are not too slow). The statistics for this problem are shown in Table 2(b).

The application of the distillation process to this problem took five seconds, and produced a distilled set with just four test cases, shown in Table 2(c).

Anagrams (P71916). This problems asks for a program that reads n pairs of lines and tells, for every pair, if the lines are anagrams of each other (i.e., if one line can be obtained from the other simply reordering the letters and forgetting the spaces and the punctuation symbols). It has the peculiarity that the number n is given at the beginning of the input. Its public test cases are shown in Table 3(a), and its statistics are shown in Table 3(b).

This problem uses the predefined *split by lines* output scheme, but requires an ad-hoc input scheme. We need to specify how to parse the input tests (for this case, the easiest is to forget the first line and group the remaining lines in pairs) and how to rewrite them (print the number of pairs followed by the pairs themselves). The following simple code is all we need to do this customization:

Table 3: Results for problem “Anagrams”.

(a) Public samples

Sample input
```
3
ROME.
MORE.
AVE MARIA, GRATIA PLENA, DOMINUS TECUM.
VIRGO SERENA, PIA, MUNDA ET IMMACULATA.
ARMY.
MAY.
```

Sample output
```
YES
YES
NO
```

(b) Statistics

Submissions	2454
Accepted	922
Rejected	1532
Compilation Error	274
Wrong Answer	949
Execution Error	309
Users with accepted problem	636
Users with rejected problem	136
Good submissions	922
Bad submissions	1093
Horrible submissions	439

(c) Distilled input test cases

```
6
R.
RRR.
.
RR.
,,,.
,.
A.
.
WINSTON CHURCHILL.
I'LL CRUNCH THIS NOW.
.
.
```

```
class CustomInputSplitter(Splitter):
    def read (self, f):
        return group(f.readlines()[1:], 2)
    def write (self, inputs, f):
        print >>f, len(inputs)
        for input in inputs:
            print >>f, input[0] + "XXn" + input[1]
```

The generated distilled file for this problem is shown in Table 3(c) and evidences some corner cases.

Brackets and Parentheses (P96529). We finally consider a problem that is more difficult from an algorithmic point of view: To read words made up of brackets and parentheses and, for each one, tell whether the brackets and the parentheses close correctly. Its public test cases are shown in Table 4(a).

Table 4: Results for problem “Brackets and parentheses”.

(a) Public samples

Sample input	Sample output
`[](([]()) `	yes
`((([])() `	no
`[](`	no
`((([])) `	yes

(b) Distilled input test cases

```
(
[]
][
]]
)]
[[[
([)[]
[((())]]
[[(())](
[(()[])][([]])(())]
```

In this case, the distillation process produced 11 individual test cases out of the 1766 private test cases. However, the distilled set has a big issue: its total size is quite big (18360 characters). Clearly, although these test cases catch all bad submissions, they are too long to be helpful to humans. Our solution to this trouble consisted in modifying the ILP so as to minimize the total size of the distilled set by weighting each variable in the objective function by the length of the corresponding input. Doing so, we obtained 13 test cases, whose total size is 1254 characters, and where three test cases still have more than 350 characters each. By removing these three long test cases, we got the distilled set shown in Table 3(b), which still catches 91% of the bad submissions.

5 CONCLUSIONS AND FUTURE WORK

In this paper we have considered how to identify small and relevant test cases for problems in online judges. The main idea of our solution is to data mine past wrong (but not too wrong) submissions to automatically extract the relevant test cases.

To do so, we have provided an elegant solution that we have applied and tested on several problems of Jutge.org, a virtual learning environment for computer programming. The results suggest that our solution is general, efficient, and generates high quality

test cases. Taking into account the amount of effort spent by problem setters in creating huge and exhaustive private test cases, it is almost deceiving to see *a posteriori* that only a few of these test cases are, in fact, relevant.

These relevant test cases can improve the feedback of educative online judges which, until now, just report a negative verdict in the event of a wrong submission. However, an important question that remains to be addressed is how to use these distilled sets in the workflow of educative online judges.

A first option would be to disclose the distilled test cases to students that request them after they have submitted a bad submission. A similar approach would be to disclose only one distilled test case for which the student's code fails (assuming there is at least one, which should happen most of the times). Although this could be a way to help beginners to spot their errors and fix them, it also could induce a "trial and error" behavior similar to the one already criticized in the introuction of this paper. At the very least, the counterexample where the code fails, being part of the distilled set, should provide a deeper insight to the users than just a random counterexample.

An alternative to mitigate the above option would be to keep distilled test cases private from students, but available to instructors. Looking both to the code of their mentored students and to the distilled test cases (where the code fails), instructors could be able to provide better individual answers, including, if deemed suitable, some of the distilled test cases.

The community of instructors using Jutge.org is currently debating these options and some variations of them.

ACKNOWLEDGEMENTS

We thank all the users of Jutge.org: without their numerous mistakes, this work would have never seen the light.

REFERENCES

Alon, N., Moshkovitz, D., and Safra, S. (2006). Algorithmic construction of sets for *k*-restrictions. *ACM Trans. Algorithms*, 2(2):153–177.

Cheang, B., Kurnia, A., Lim, A., and Oon, W.-C. (2003). On automated grading of programming assignments in an academic institution. *Computers & Education*, 41(2):121–131.

Chvatal, V. (1979). A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, 4(3):233–235.

Douce, C., Livingstone, D., and Orwell, J. (2005). Automatic test-based assessment of programming: A review. *ACM Journal on Educational Resources in Computing*, 5(3).

Forišek, M. (2006). Security of Programming Contest Systems. In Dagiene, V. and Mittermeir, R., editors, *Information Technologies at School*, pages 553–563.

Giménez, O., Petit, J., and Roura, S. (2009). A pure problem-oriented approach for a CS1 course. In Hermann, C., Lauer, T., Ottmann, T., and Welte, M., editors, *Proc. of the Informatics Education Europe IV*, pages 185–192.

Gurobi (2013). Gurobi Optimizer Reference Manual. `http://www.gurobi.com`.

Ihantola, P., Ahoniemi, T., Karavirta, V., and Seppälä, O. (2010). Review of recent systems for automatic assessment of programming assignments. In *Procs. of the 10th Koli Calling International Conference on Computing Education Research*, pages 86–93. ACM.

Joy, M., Griffiths, N., and Boyatt, R. (2005). The BOSS online submission and assessment system. *ACM Journal on Educational Resources in Computing*, 5(3).

Karp, R. M. (1972). Reducibility Among Combinatorial Problems. In Miller, R. E. and Thatcher, J. W., editors, *Complexity of Computer Computations*, pages 85–103.

Kurnia, A., Lim, A., and Cheang, B. (2001). Online judge. *Computers & Education*, pages 299–315.

Leal, J. P. and Silva, F. (2003). Mooshak: a web-based multi-site programming contest system. *Software: Practice and Experience*, 33(6):567–581.

Petit, J., Giménez, O., and Roura, S. (2012). Jutge.org: an educational programming judge. In *Procs. of the 43rd ACM technical symposium on Computer Science Education*, SIGCSE '12, pages 445–450. ACM.

Singh, R., Gulwani, S., and Solar-Lezama, A. (2013). Automated feedback generation for introductory programming assignments. In Boehm, H.-J. and Flanagan, C., editors, *PLDI*, pages 15–26. ACM.

Tonin, N., Zanin, F., and Bez, J. (2012). Enhancing traditional algorithms classes using URI online judge. In *2012 International Conference on e-Learning and e-Technologies in Education*, pages 110–113.

Verdú, E., Regueras, L. M., Verdú, M. J., Leal, J. P., de Castro, J. P., and Queirós, R. (2012). A distributed system for learning programming on-line. *Computers & Education*, 58(1):1 – 10.

A Technology Enhanced Assessment System for Skill and Knowledge Learning

Enosha Hettiarachchi[1], Enric Mor[2], Maria Antonia Huertas[2] and M. Elena Rodriguez[2]
[1]*Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Spain*
[2]*Computer Science, Multimedia and Telecommunication Studies, Universitat Oberta de Catalunya, Barcelona, Spain*
{khettiarachchi, emor, mhuertass, mrodriguezgo}@uoc.edu

Keywords: e-Assessment, Higher-order Skills Acquisition, Knowledge Acquisition, Interactivity, Online Higher Education, Technology-enhanced Assessment.

Abstract: This paper presents a technology-enhanced assessment system that can be used for both skill and knowledge learning. For this purpose, a general technology-enhanced assessment system is designed and developed for an online logic course at a fully online university, taking into account e-learning standards and specifications, which can be easily adapted to any institute and subject requiring a high level of skill learning. Through this system, both learning and formative assessment facilities are provided to students. To evaluate its effects on student learning process, the system is applied in a real logic course at the Universitat Oberta de Catalunya (UOC). Based on the evaluation, it shows that students' are more engaged with the system and, as a result, their performance in the subject had improved. Also, based on the feedback obtained through the evaluation, it shows that students are satisfied with the facilities and assessments provided by the system. Overall, the introduction of the technology-enhanced assessment system for skill and knowledge learning has yielded some interesting results.

1 INTRODUCTION

Nowadays, e-learning systems have become a common form of media for providing education transparency in society (Krishnamurthy & O'Connor, 2013). E-Learning systems that provide higher interactivity and feedback makes students motivated and engaged to learn more efficiently (Bull & Mckenna, 2004; Sadler, 2013). As a result, new challenges have arisen for educators and technologists and, among them, providing student engagement and assessment can be noted.

Technology-enhanced assessment, also known as e-assessment, deals with methodologies, tools and processes where information and communication technologies are used for the delivery of assessment activities, the recording of responses and the provision of feedback (Cook & Jenkins, 2010; Daly et al., 2010; JISC, 2007). Traditionally, e-assessment has been used for testing the acquisition of declarative knowledge where students are required to select a predetermined response based on factual recall like, for example, the simple multiple-choice question types (Bull & Mckenna, 2004; de Bruyn et al., 2011). However, cognitive skills where students have to apply their analytic, creative and constructive skills cannot be assessed via multiple-choice tests and equivalent forms (Gruttmann et al., 2008; Majchrzak & Usener, 2011). As Crisp (2009, 2010) stated, in order to test higher order capabilities, it is needed to design sophisticated assessment tasks, but the difficulty and workload in designing such tasks are considerable. Skill e-assessment when offered is, usually, subject dependent and technologically complex because of the computational difficulties to represent and simulate higher order cognitive questions and its automatic marking. Hence, most of the technology-enhanced assessment tools that are currently available are either developed specifically for the particular subject content or only offer a simple type of questions that can be used only for the assessment of knowledge acquisition. Therefore, one of the main disadvantages is that there is no general type of tool that can be used for assessment of both skill and knowledge activities (Hettiarachchi, et.al, 2013)

Another challenge is associated with the software used for e-assessment systems. According to Bull & Mckenna (2004), several issues have been identified as critical when it comes to the decision-making process. They can be noted as interoperability,

integration with existing systems, scalability, performance level, limitations associated with upgrading procedures, support and maintenance and security and accessibility.

Considering the above, this paper introduces a technology-enhanced assessment system, which goes beyond just the combination of existing assessment systems, for both skill and knowledge learning and assessment in an online educational environment. In addition to that, the system is designed and developed as a standardized and a general system that can be easily adapted by any domain or subject. The goal of this paper is to evaluate whether the developed e-assessment system works correctly and this was evaluated through a real online Logic course at a fully online university.

The rest of the paper is organized as follows. Section 2 presents a general introduction to the skill and knowledge assessment with some of the common e-assessment tools and systems. Section 3 explains the proposed technology-enhanced assessment system. Section 4 presents the data analysis and results based on the evaluation of the system. Finally, the paper ends with the conclusions.

2 SKILL AND KNOWLEDGE E-ASSESSMENT

Assessment activities can be divided broadly into two types such as skill and knowledge assessment (Crisp, 2009). Knowledge can be specified as the recall or recognition of specific items. It can be more elaborate as remembering of previously learned materials and contents (Bull & Mckenna, 2004; de Bruyn et al., 2011). This may involve the recall of a wide range of content, from specific facts to complete theories, but all that is required is the bringing to mind of the appropriate information. Knowledge e-assessment mostly uses simple forms of questions such as Multiple Choice Questions (MCQ), multiple responses, short answers and fill-in the blanks. They are generally easier to mark both as automatic and human means. This type of assessment is quicker in delivery, gives more specific and directed feedback to individuals and can also provide greater curricular coverage (McAlpine, 2002). At the same time, they can be limited in scope and can occasionally degenerate into a 'quiz' of facts about the area of study.

Skill can be defined literally as a practiced ability, expertness, technique, craft and art. Higher-order cognitive skills are typically required for solving exercises encountered in the natural sciences, including computer science and mathematics (Gibbs & Simpson, 2004). Skill e-assessment activities are often associated with a constructivist view of learning and it is best suited when there may be a difference of opinion based on interpretation (Crisp, 2007) or to assess higher cognitive skills. However, they can be time consuming to set and mark. They also require greater marking proficiency than knowledge assessment activities, involving training markers or detailing criteria (McAlpine, 2002).

2.1 e-Assessment Tools and Systems

The main characteristics of an e-assessment system have been widely studied. The most important are (Bull & Mckenna, 2004; Sitthiworachart et al., 2008; Tselonis & Sargeant, 2007):

- monitoring student progress through frequent assessments,
- applying a variety of interactive question types and promoting student engagement,
- automatic marking, weighted-average grade calculation and immediate feedback,
- supporting flexible and adaptive learning, and personalization of assessment activities,
- monitoring question quality using statistical analysis and
- reducing the potential for cheating by randomizing questions along with timers, and sharing questions via question banks.

Some universities and educational institutes offer e-assessments; but they are mostly based on knowledge assessment rather than skill assessment activities (Marriott, 2009; Pachler et al., 2010). One of the reasons is that most of the tools support only simple type of the questions such as MCQ. At the same time, tools that are based on skill assessment activities usually depend on a specific subject context because of their complex semantics.

Since this research is focused on a logic course, the e-learning and e-assessment tools used for logic were analysed. Currently, there is a quite large sample of tools used for learning logic courses, many of them can be categorized as Intelligent Tutoring Systems (ITS); but not so much for e-assessment. The main characteristic of an ITS for learning is providing customized assistance and feedback to students while simulating the presence of an e-tutor, however, they lack many of the characteristics of an e-assessment system. There is an extensive discussion on e-assessment tools in Crisp (2007) and on ITS for teaching logic in

(Huertas, 2011). In the case of logic, many tools fall into the category of ITS, for example: Pandora (Imperial College London, 2013), Organon (Dostalova & Lang, 2007), and AELL (Huertas, 2011), but none was fulfilling all the general features of an e-assessment system, noted above.

3 TECHNOLOGY-ENHNACED ASSESSMENT (TEA) SYSTEM

The context of this research is a first year Logic course of a Computer Science degree in the Universitat Oberta de Catalunya [www.uoc.edu]. Logic is a good case study for this research because it requires a high level of skill and knowledge learning. In this course, an ITS for learning logic, AELL, was developed and used previously, with the aim of providing a tool to facilitate the learning process (Huertas, et.al, 2011). The tool assisted students in different kind of activities by guiding and informing them of the correctness of their solutions. The main aim of the ITS tool was to provide learners with more practice through automatically graded exercises, for learning purposes. For assessment, students had, as is traditional, a set of assessment activities, provided through the ITS, that were the same for all the students. Therefore, they had the possibility to copy answers from each other.

In order to provide a fully formative e-assessment experience, as mentioned before, it was needed to introduce both practice and assessment of skill and knowledge acquisition in order to motivate students and provide a rich e-assessment experience while minimizing copying and cheating. In particular, main characteristics of e-assessments, not present in the ITS tool used for practice, had to be introduced. Therefore, we decided to go a step forward the use of the ITS and design a new system to provide e-assessment and feedback in an interactive way.

The proposed Technology-Enhanced Assessment (TEA) system was designed and developed iteratively following a user centered design approach (Bevan, 2003). To identify the features and functionalities of the system, data were collected in the form of surveys, focus groups and interviews. Also, interfaces were tested and prototypes were built in order to evaluate the system and make it match teacher's goals and student's learning and assessment needs.

The TEA system was designed to provide both practice and assessment in both skill and knowledge acquisition. For practice, students were provided with facilities such as learning materials and practice tests. For assessment, the system provided assessment tests. Both practice and assessment tests included interactive feedback and based on the feedback students were able to attempt the tests till they obtain the required marks needed to master the knowledge and skills required. For assessment to be effective, feedback must not only be provided, but also understood by students and acted on in a timely fashion (Jordan, 2009). Therefore, the feedback provided through the system was immediate and detailed with guidance. Based on that, students should be able to interactively learn their errors or mistakes and obtain a higher mark in the subsequent attempt. For assessment test, some restrictions were imposed with time and attempts to motivate students as well as to offer individual questions and the assessment atmosphere.

3.1 Architecture of the TEA System

In this section, the architecture for a general TEA system, from which the case of the TEA system for the Logic course is developed, is presented. It consists of five modules: skill assessment module, knowledge assessment module, progress bar, competencies and gradebook.

Skill assessment module provides dynamic and interactive questions for both practice and assessment tests where students have to construct the answers with the guidance of feedback, errors and hints. In the case of the TEA for the Logic course, this module uses an ITS for Logic practice. In the case of a different subject, a tool of the type of an ITS can be used. Knowledge assessment module also provides both practice and assessment tests with simple knowledge type of questions such as MCQ. Also, for these questions, feedback is provided for each step performed by the student. Progress bar is a module that provides visual guidance for helping students to understand their progress with respect to the course. It shows the total progress obtained by each student along with the graphical presentation of activities completed, to be completed and not completed. Competencies module allows teachers to understand the competencies achieved by students in a particular course. These competencies are selected based on the marks obtained by students for a particular activity or test. Students can view the competencies they have achieved as a progress bar and a list of tables. Gradebook module is used to display grades and outcomes obtained by students for each activity or test. These components help

teachers to track students learning progress throughout the whole course period.

Out of the five modules mentioned, progress bar, competencies and gradebook are taken as the Basic TEA System as these are the basic functionalities of the main general TEA system and they are not subject dependent. In addition to that, the Basic TEA system is capable of storing log data related to students' participation in the activities and statistics.

Both the knowledge assessment and skill assessment modules are independent modules, where the skill assessment module is, usually, subject dependent. They are connected with the Basic TEA system using a plug-in, developed for this research purpose.

As shown in Figure 1, the users log into the LMS of the educational institution and automatically navigate to the TEA system through the single sign-on facility provided by the IMS Learning Tools Interoperability (LTI) specification (IMS GLC, 2013). The principal concept of LTI is to establish a standard way of integrating rich learning applications with platforms like learning management systems or other educational environments. Also, skill and knowledge assessment modules are linked with the Basic TEA system with the aid of the developed plug-in. For transferring data such as user data, grades, time spent and attempts, from the two modules to the Basic TEA system and from the TEA system to the LMS, OAuth protocol (OAuth, 2013) is used together with the IMS LTI specification. This protocol is used to secure its message interactions between the tools. The connection and the communication between tools are carried-out through both message-based and service-based connections (Hettiarachchi et al., 2012).

In the case of the Logic course, an MCQ module was used as the knowledge assessment and our previously developed Logic ITS tool, enhanced into an assessment tool with the features of incorporating a large database of questions based on different difficulty levels, randomized selection of questions, immediate and detailed feedback, limited time and limited attempts, was used for the skill assessment. Since both the skill and knowledge assessment modules are independent modules, instead of these tools, any other tool can be taken and easily connected with the Basic TEA system using the developed plug-in in a secure and interoperable manner. Therefore, depending on the context, any other tool can be used.

4 EVALUATION

The goal of this paper is to evaluate whether the developed e-assessment system works correctly. Also, the impact of the system in students' performance and engagement in the classroom was evaluated.

To evaluate the proposed TEA system and analyse its impact on students, the system was used in a Logic course of the Universitat Oberta de Catalunya. The course duration was 14 weeks with the participation of 38 students.

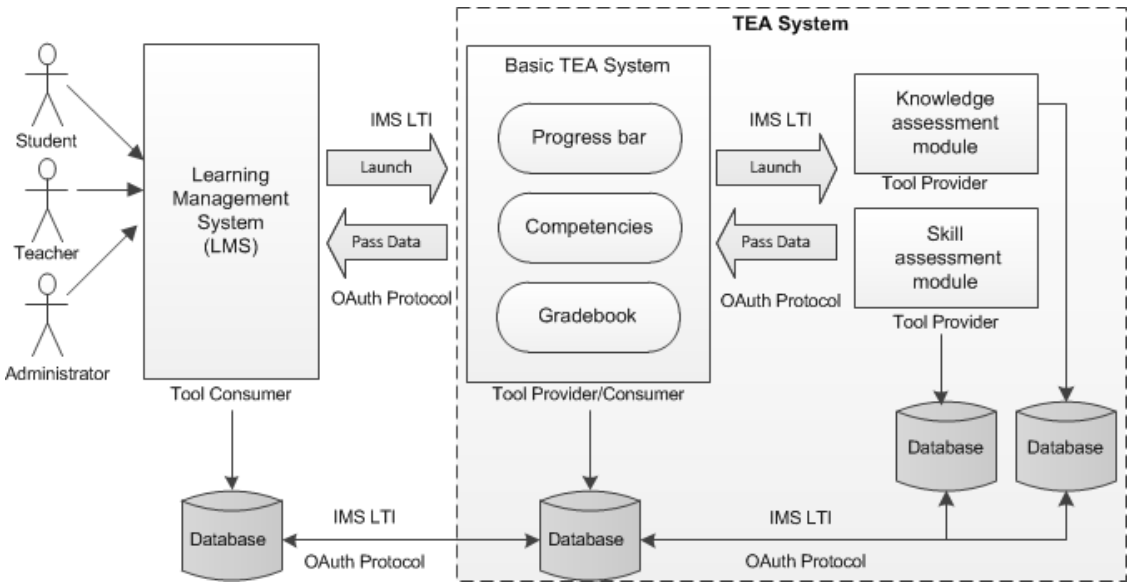


Figure 1: Architecture of the system with main components of the TEA system.

For the evaluation, a formative assessment model was introduced into the Logic course for both skill and knowledge assessment. For the formative assessment, students were provided with both practice and assessment activities through the system. To motivate students to practice with the interactive questions and feedback, a restriction was imposed as students needed to obtain a minimum pass mark in the practice activities to move to assessment activities. Questions within the assessment activities were selected randomly from a large question bank to minimize cheating. Also, soon after completion of a particular test, students were offered with immediate detailed feedback. Also, students were given a face-to-face 2 hour final examination. The final grade of the course comprised of 35% of marks in the formative assessment and 65% of marks in the summative assessment.

4.1 Analysis and Results

Data were collected mainly from two sources. On the one hand, data related to student engagement with the TEA system were obtained through the system log. On the other hand, a questionnaire was given to the students to obtain their feedback regarding the learning experience with the TEA system. This questionnaire was also used to obtain students' perceptions about the system, the improvements needed to be carried-out in the future, and also to draw conclusions regarding the student experience with the system. The questionnaire consisted of 28 questions consisting of open-ended, yes/no and five-point Likert scale questions. These questions were divided into four sections such as learner information, student satisfaction, formative assessment and assessment model.

To analyse students' engagement, data were obtained from the system logs about student participation using the TEA system. The results of the analysis are shown in Figure 2. According to that, each student accessed the system minimum 5 times during a particular day. The TEA system consisted of session time-out duration and therefore, students might have had to login to the system more than once during the day. This could be shown as the highest peaks in the diagram. The majority of these peaks had occurred when it was closed the deadlines of the assessment activities (AA corresponds to Assessment Activities).

High peaks at the beginning shows that, students have used the TEA system more to get familiar with its features. Overall, students have continuously used the TEA system throughout the whole duration of the course. This could be due to the fact that students had appreciated the facilities, such as interactivity, immediate feedback and marks, provided by the TEA system for practice purposes.

To explore students learning experience, a set of questions was introduced into the questionnaire.

Regarding the student satisfaction, four questions were given to the students. About the instructions presented for answering the questions, 89% of students agreed by answering they were presented in

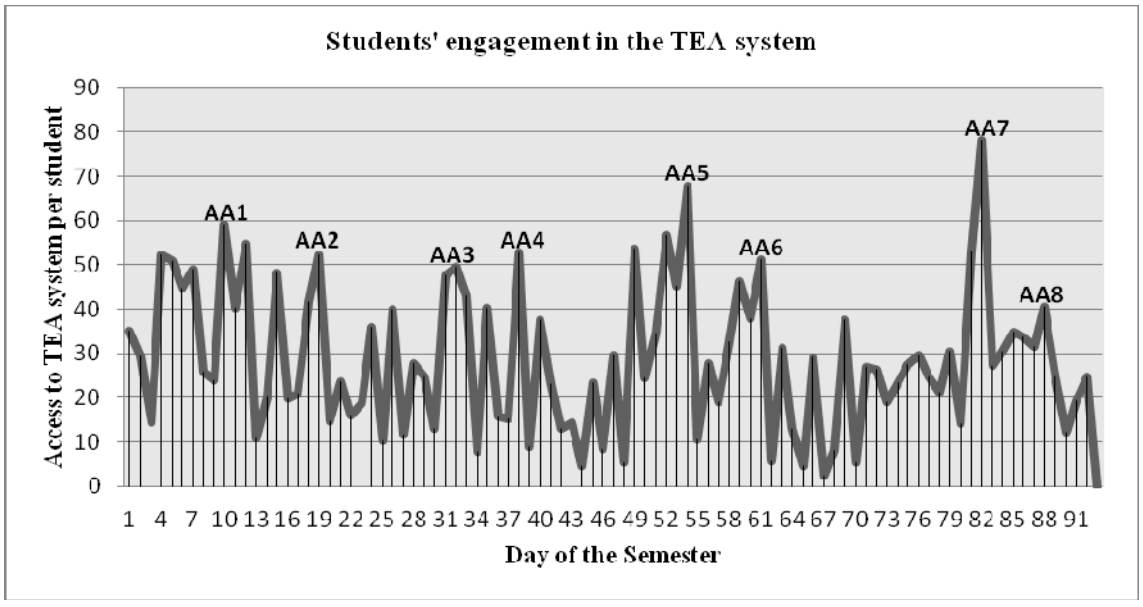


Figure 2: Students' engagement in the TEA system.

a clear and concise manner. Also, 68% of the students agreed that the automatic grades offered through the system were very good. At the same time, 89% of students were satisfied with the questions provided for both practice and assessment. Overall, students were satisfied with the TEA system. For formative assessment, students' opinions about practice, assessment, feedback and their relationship to improving learning process were evaluated. First, it was needed to understand whether it was helpful to practice before attempting the assessments and 74% of students agreed and they further mentioned that practicing using the system helped them to evaluate the skills and knowledge acquired as well as they were able to practice and get a comprehensive review of the questions offered in the assessments. When it comes to the automatic feedback, 89% of students agreed that feedback provided by the TEA system was satisfactory. This can be taken as a reason due to detailed immediate feedback, hints and suggestions introduced in the system. Based on the marks offered by the TEA system, 89% students agreed that the marks fit their knowledge and skills developed. Therefore, it can be stated that the TEA system was capable of offering correct marks or grades to fit the skills and knowledge acquired by students. Also, 79% of students considered that practice and assessment tests provided were helpful for learning skills related to the course. Furthermore, students also agreed that both tests helped them to understand the topics covered in the materials. Therefore, to find the difficulty of the assessments, when asked about the average number of attempts students had to complete in-order to achieve the minimum score, 74% mentioned 2 attempts and another 11% mentioned 3 attempts. As a conclusion an average of 2 attempts were needed to obtain the minimum score. At the same time, it can be taken as an indicator that the assessments were of medium difficulty level and they are suitable for assessment of knowledge and skills. Also, 79% of students strongly agreed that assessment tests have helped them to evaluate their strengths and weaknesses in the Logic course. To get students' opinions about the use of assessment tests in the subject, whether they would have learned the same if they did not have assessment tests and 89% of students answered by saying no. Therefore, it shows that students have valued the importance of assessments in the learning process. When it comes to evaluating the progress of doing tests using the progress bar, only 74% agreed, whereas some students have mentioned, it was useful but not essential. About the usefulness of the competency module, 79% of students agreed by saying it was useful, whereas the rest of the 21% did not agree. When asked about the reasons most of them have mentioned that they have not seen the module since they have not been informed about it. However, when asked about grades and outcomes, interesting 100% agreed that both grades and outcomes were useful information. Finally, an open-ended question was given to obtain students comment and suggestions about the system. Overall, students liked the system, unless some students have mentioned that the time given, 2 hours, for the assessment test was not enough.

5 CONCLUSIONS

The TEA system was introduced to support the student learning process based on both skill and knowledge acquisition. The system was designed to offer interactivity in e-assessments with an architecture that favours its application to different domains and its connection with different existing LMS in a secure and interoperable way.

This system provided both practice and assessment facilities for students to improve their learning process. Therefore, students were constantly engaged in the system and as a result, their performance in the formative assessment and summative assessment had improved. Also, the information provided through the progress bar and competencies module helped students to evaluate their own progress. However, most of the students have not fully utilized these features, as they were not informed about them. Therefore, in the future, students have to be informed at the beginning of the course about the facilities offered through the progress bar and the competencies module.

Student participation data in the TEA system showed that students were constantly engaged in the system for both practice and assessment purposes. It also showed that students were more engaged in the system when it was close to a completion date of an assessment test. Also, students had used the system even after the completion dates of the assessments. This showed that students had used the system to prepare for the final examination. Overall, as a conclusion, it can be stated that students had constantly engaged in the TEA system throughout the course duration. Overall, students were satisfied with the TEA system, formative assessment, assessment model, course scheduling, marks and feedback provided. Students were also satisfied with the detailed and immediate feedback and they

believe that doing practice activities had helped them to perform better in the assessments, and to evaluate the skill and knowledge acquired. Also, according to students, both practice and assessment tests helped them to evaluate their strength and weakness in the Logic subject, and learn skills related to the subject. However, some students mentioned that it was a bit stressful and the allocated time was not enough to complete some of the questions related to skills. Therefore, as improvements it is needed to consider about the time given for the assessments, mostly in the sections where students have to construct the answer using the skill assessment module. At the same time, a complete schedule with assessments has to be displayed in the main course page.

Although this research was carried out in a fully online environment, the developed TEA system with the formative assessment approach based on skills and knowledge can be extended to blended courses. In the future, for testing the interoperability of the TEA system, it will need to be introduced into other courses based on skills and knowledge as well as by connecting with other LMS.

As a general summary, the technology-enhanced assessment system was capable of supporting students learning process and as a result students' performances in the online classroom had improved.

ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Ministry of Science and Innovation funded Project MAVSEL (ref. TIN2010-21715-C02-02), Computer Science, Multimedia and Telecommunication Studies Department of the Universitat Oberta de Catalunya (UOC) and the Internet Interdisciplinary Institute (IN3) of the UOC.

REFERENCES

Bevan, N., 2003. UsabilityNet Methods for User Centred Design. Human-Computer Interaction: Theory and Practice (Part 1), *Proceedings of HCI International,* 1, 434 - 438.

Bull, J., & Mckenna, C., 2004. *Blueprint for Computer-Assisted Assessment (Vol. 2)*. RoutledgeFalmer, London.

Cook, J., & Jenkins, V., 2010. *Getting Started with E-Assessment*. Bath. Available at: <http://opus.bath.ac.uk/17712/> [Accessed on 31 January 2014]

Crisp, G., 2007. *The e-Assessment Handbook*. Continuum International Publishing Group, London.

Crisp, G., 2009. *Interactive E-Assessment: Moving Beyond Multiple-Choice Questions*. Centre for Learning and Professional Development. Adelaide: University of Adelaide.

Crisp, G., 2010. Interactive E-Assessment - Practical Approaches to Constructing More Sophisticated Online Tasks. *Journal of Learning Design, 3* (3), 1 - 10.

Daly, C., Pachler, N., Mor, Y., & Mellar, H., 2010. Exploring Formative E-Assessment: Using Case Stories and Design Patterns. *Assessment & Evaluation in Higher Education, 35* (5), 619 - 636.

de Bruyn, E., Mostert, E., & Schoor, A., 2011. Computer-based Testing - The Ideal Tool to Assess on the Different Levels of Bloom's Taxonomy. *In 14th International Conference on Interactive Collaborative Learning (ICL2011),* 444 - 449.

Dostalova, L., & Lang, J., 2007. ORGANON The Web Tutor for Basic Logic Courses. *Logic Journal of IGPL, 15* (4), 305 - 311.

Gibbs, G., & Simpson, C., 2004. Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education, 1* (1), 3 - 31.

Gruttmann, S., Böhm, D., & Kuchen, H., 2008. E-assessment of Mathematical Proofs: Chances and Challenges for Students and Tutors. *International Conference on Computer Science and Software Engineering (IEEE),* 612 - 615.

Hettiarachchi, E., Huertas, M. & Pera, E., 2013. “Skill and Knowledge E-Assessment: A Review of the State of the Art” [online working paper]. (Working Paper Series; WP00-000). IN3 Working Paper Series. IN3 (UOC). In press.

Hettiarachchi, E., Huertas, M. A., Pera, E. M., & Guerrero-Roldan, A. E., 2012. An Architecture for Technology-Enhanced Assessment of High Level Skill Practice. In *2012 IEEE 12th International Conference on Advanced Learning Technologies (ICALT)*, 38-39. IEEE.

Huertas, A., 2011. Ten Years of Computer-based Tutors for Teaching Mathematical Logic 2000-2010: Lessons Learned. In *P. Blackburn et al. (Ed.): Third International Congress on Tools for Teaching Mathematical logic (TICTTL 2011), LNAI 6680,* 131 - 140. Springer, Heidelberg.

Huertas, A., Humet, J. M., López, L., & Mor, E., 2011. The SELL Project: A Learning Tool for E-Learning Logic. *Learning and Instruction, 6680/2011*, 123 -130.

Imperial College London, 2013. *Pandora IV*. Available at: <http://www.doc.ic.ac.uk/pandora/newpandora/index.html> [Accessed on 31 January 2014]

IMS GLC, 2013. *Learning Tools Interoperability*. Available at: <http://www.imsproject.org/lti> [Accessed on 31 January 2014]

JISC, 2007. *Effective Practice with e-Assessment: An Overview of Technologies, Policies and Practice in Further and Higher Education*. Available at: <http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf> [Accessed on 31 January 2014]

Jordan, S., 2009. Assessment for Learning : Pushing the

Boundaries of Computer-Based Assessment. *Practitioner Research in Higher Education*, *3* (1), 11 - 19.

Krishnamurthy, A. and O'Connor, R., 2013. An Analysis of the Software Development Processes of Open Source E-Learning Systems, *Systems, Software and Services Process Improvement*. Springer Berlin Heidelberg, 60-71.

Majchrzak, T. A., & Usener, C. A., 2011. Evaluating the Synergies of Integrating E-Assessment and Software Testing. *In Proceedings of Information Systems Development Conference (ISD2011)*. Springer, 2011.

Marriott, P., 2009. Students' Evaluation of the use of Online Summative Assessment on an Undergraduate Financial Accounting Module. *British Journal of Educational Technology, 40* (2), 237 - 254.

McAlpine, M., 2002. *Principles of assessment*. CAA Centre, University of Luton.

OAuth, 2013. *OAuth - Getting Started*. Available at: <http://oauth.net/documentation/getting-started/> [Accessed on 31 January 2014]

Pachler, N., Daly, C., Mor, Y., & Mellar, H., 2010. Formative E-Assessment: Practitioner Cases. *Computers & Education, 54* (3), 715 - 721.

Sadler, D. R. (2013). Opening up feedback. Reconceptualising *Feedback in Higher Education: Developing Dialogue with Students*, 54.

Sitthiworachart, J., Joy, M., & Sutinen, E., 2008. Success Factors for E-Assessment in Computer Science Education. In *C. Bonk et al. (Eds.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, 2287 - 2293.

Tselonis, C., & Sargeant, J., 2007. Domain-specific Formative Feedback through Domain-independent Diagram Matching. In *F. Khandia (Ed.), 11th CAA International Computer Assisted Assessment Conference,* 403 - 420. Loughborough, UK: Loughborough University.

Developing Web-based Tutorial Modules to Support Training for Group-based Mentoring

Sharon Locke[1], Georgia Bracey[1] and Stephen Marlette[2]
[1]*Center for STEM Research, Education, and Outreach, Southern Illinois University Edwardsville, PO Box 2224, Edwardsville, Illinois, U.S.A.*
[2]*Department of Curriculum & Instruction, Southern Illinois University Edwardsville, Edwardsville, Illinois, U.S.A.*
{slocke, gbracey, smarlet}@siue.edu

Keywords: Mentoring, Blended Learning, Web-based Learning, Self-Efficacy, STEM, Robotics.

Abstract: This paper describes the development and implementation of a series of interactive online modules that serve as the core component of a mentor training program for middle school teachers in the Botball® Educational Robotics Program. Botball is an international team-based robotics competition for secondary students. Over the course of a season, students work together to design, build, and program a robot that will meet a particular challenge. Teachers play a critical role in this process as team mentors, but there is currently no mentor training available that is easily accessible to teachers in a variety of geographic locations. To meet this need, and as part of a larger research study, three STEM educators at a US Midwestern university created five Web-based modules designed to introduce users to the fundamentals of group mentoring. These research-based tutorials provide on-demand distance training that is engaging and reflective. We will discuss formative evaluation of the training gleaned from responses to open-ended survey questions given to the initial group of 33 teachers who piloted the modules as well as from interactive webinars offered throughout the training.

1 INTRODUCTION

Mentoring plays a critical role in youth STEM (science, technology, engineering, and mathematics) activities. The relationship between student and mentor has been shown to have an impact on the success of the activity as well as the attitudes, interests, and learning of the participants. Moreover, students with positive perceptions regarding their mentor's effectiveness tend to have increased positive perceptions towards STEM careers (Weinberg et al., 2007). It is clear that adults who are effective mentors are more likely to produce positive outcomes from their students. Unfortunately, most adults do not have specific training in the practice of mentoring, and there are few training programs available that are easily accessible to those who may live miles away from the nearest university or training center.

As part of a university study examining components of effective mentor training for after-school robotics programs, five Web-based tutorials were developed as one piece of a larger training program. The larger program consisted of interactive webinars, a daylong face-to-face meeting, and the Web-based modules. As study participants moved through the program, they had opportunities to engage in activities with university faculty and staff to reinforce and reflect on the content presented in the modules. While many of these activities focused on mentoring in the context of Botball, the modules were designed to be tutorials that could stand on their own and be useful in any group-mentoring situation.

2 THEORETICAL FRAMEWORK

Two principles guided the content development of the Web-based modules. First, general best practices in mentoring--taken from theory and from empirical evidence--formed the basis of two of the five modules. These modules focused on team-building theory (Tuckman, 1965) plus practical issues such as formation of team members' roles, goal-setting, and creating cohesion. The content for a second pair of modules was developed around the theory of self-efficacy and its four components: mastery

experience, vicarious experience, social persuasion, and physiological reaction (Bandura, 1977, 1997).

Pedagogy for the modules took a constructivist approach. Constructivism states that learners "learn actively and construct new knowledge based on their prior knowledge" (Huang, 2002). With this in mind, the modules were designed to be interactive and to encourage the learners to connect the content to their own experience and understanding. Thoughtful, open-response questions throughout the modules provided opportunity to interact with and reflect on the content.

3 MODULE DESIGN

The purpose of the Web-based modules was to present concepts of mentoring best practices and mentoring for self-efficacy in a self-paced format. The modules served as an introduction to mentoring that was later reinforced through a face-to-face workshop and webinars. Since mentors would complete the modules individually, the design incorporated engaging elements such as video clips and reflective questions that supplemented a narrative text. Summative questions at the end of the module checked participants' understanding of the content.

The team selected the Moodle development platform because it has an intuitive interface and is free to use. An outside consultant familiar with Moodle transferred the content into Moodle and developed the navigation features. The project team provided feedback on the graphic design and reviewed the online version of the module for accuracy.

3.1 Multimedia Presentation

The module narrative integrated images and videos to enhance interest and provide opportunities for participants to apply their knowledge (Figure 1). Video clips and images were selected from freely available content on the Internet and imported into the module platform. Reflective questions were embedded throughout the module.

After a concept was introduced using narrative or a video clip, participants responded to a prompt in an open form field (Figure 2). The reflective questions asked participants to draw on their own previous experiences in mentoring and often required application of new concepts presented in the module.

At the end of each module participants also completed a set of summative questions that served as a review of the material (Figure 3).

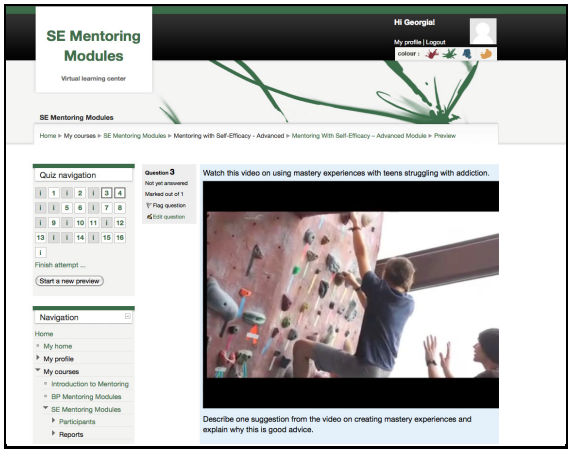


Figure 1: Module video interface.

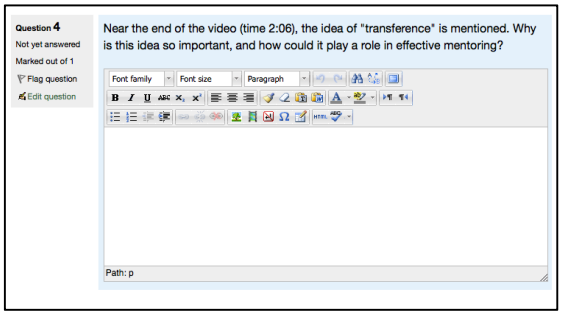


Figure 2: Example of reflective question.

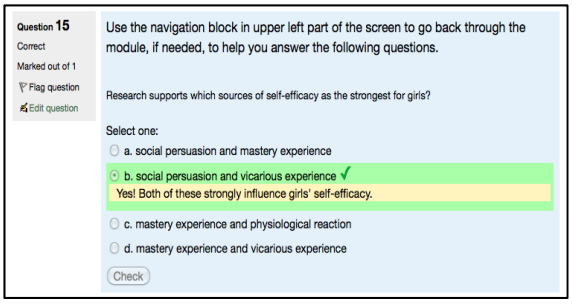


Figure 3: Example of summative question.

4 DATA COLLECTION

The modules include embedded reflective questions for participants to complete as they progress through the mentor training. The purpose of the questions is to help mentors identify and recall previous experiences with mentoring and reflect on the relevance of that prior experience to the new context. Researchers were able to access participants' answers in a database and use the

information to determine participants' prior mentoring experience, views on mentoring, and degree of engagement with the module content.

The open-ended answers provided detailed information about participants' previous mentoring experience and opinions about the mentoring process. In some cases participants were able to make meaningful connections between the module content and their own personal experiences, as shown in this response:

> *Adolescents love social interaction. The students in my class love anything that is hands-on where they can work with peers and socially interact. In a group mentoring program students must work together as a team. They get the chance to socially interact and make decisions together. They are forced to listen to each other's ideas and work together.*

The responses also served as a basis for preparing the webinar presentations, with example responses anonymously shown during the webinar. Participants were invited to explain their response in more detail if it appeared during the webinar and they were comfortable revealing their identity.

Multiple-choice questions at the end of each module are designed to reinforce the content and serve as a check for understanding; for example, the following question occurs at the end of the basic self-efficacy module: *Which one of the sources of self-efficacy is operating when a parent points out the positive and negative aspects of a child's work?* Participants receive feedback if they choose an incorrect answer and can go back to change their answer. After a participant finishes all the questions, the module is recorded as "Completed" in the system.

Participant responses to the embedded assessments in the modules also serve as data for the research study. Qualitative textual analysis of the open-ended responses can reveal themes for participants' views about mentoring that may influence how they respond to training. For example, many participants used mentoring examples from their own classroom teaching and some considered themselves to be experienced mentors. Participants who identify themselves as "experts" before the training may be less engaged with the ideas presented in the modules.

5 MODULE IMPLEMENTATION

The research project required that thirty-three middle-school teachers complete the online modules. Every teacher received a webinar introduction to the modules from the three STEM educators, and then completed the modules on their own time. The introductory presentation described the goals of the mentor training, provided the web link and login instructions for the Moodle site, and demonstrated how to navigate through the site to complete the assignment. Participants had ten days to complete their assigned modules. Technical assistance was available via email or phone.

The timing of the mentor training was planned so that mentors were adequately prepared to use the mentoring strategies to support their teams during the competition season. Mentors also participate in a Botball technical workshop seven weeks prior to the final Botball tournament. The following timeline illustrates the format, content, and delivery of mentor training:

6 weeks prior to Botball technical workshop (January): Mentors participated in a one-hour webinar that introduced the project goals, schedule, Moodle platform, and online modules.

5 weeks prior to Botball technical workshop: On their own, mentors completed online tutorials on mentoring best practices and/or self-efficacy. Available modules include an introductory module on theories, models, and benefits of mentoring; two best practices modules (basic and advanced) on research-based best practices in mentoring; and two self-efficacy modules (basic and advanced) on mentoring for self-efficacy.

4.5 weeks prior to Botball technical workshop: Mentors attended a one-hour webinar that reviewed the content presented in the online modules.

3 weeks prior to Botball technical workshop: Mentors participated in 8 hours of on-site training at their location. The workshop provided time for small-group activities and discussion and reflection on potential mentoring scenarios. The face-to-face workshop closely followed the content of the online modules.

Botball technical workshop (February): Mentors received the two-day Botball technical training, including practice with the robotics kits, design challenge, scoring, and rules and begin the mentoring process.

6 weeks prior to Botball tournament: Mentors participated in a two-hour webinar that reviewed and reinforced the key principles of mentoring (best practices and/or self-efficacy) and addressed mentor questions.

3 weeks prior to Botball tournament: Mentors participated in a two-hour webinar that addressed mentor questions or problems as related to best

practices and/or self-efficacy and reminded mentors of goals for implementation.

Week 7 of competition season (April): Botball tournament

6 MODULE EVALUATION

The developers had multiple avenues for evaluating the quality and impact of the online modules. The embedded questions provided insights into the level of engagement of the participant, e.g. through the length and quality of responses to open-ended questions. In addition, three webinars were conducted after the mentors had completed the modules. The second webinar provided a review and discussion of the module content, giving mentors a chance to ask questions and the STEM educators an opportunity to clarify concepts. The third and fourth webinars prompted mentors to reflect on how well they were able to apply the mentoring strategies presented in the modules during their robotics team meetings. All webinars were recorded, enabling researchers to transcribe and review the conversations and determine how well the module content was being transferred into practice. Lastly, a survey with open-ended questions about the mentor training was administered after the Botball robotics season had ended.

Data analysis for the preliminary evaluation focused on participants' responses to three open-ended survey questions: 1) What aspects of the mentor training did you find to be least helpful? 2) Did you encounter barriers or challenges to implementing the mentoring strategies? If so, please explain, and 3) Would you want to be a Botball mentor again? Please explain. Four members of the research team used open and analytical coding to look for emerging themes in the response text. Forty percent of respondents indicated that all of the training was useful, while less frequent themes included not enough time to complete all activities, repetition of the content across the three modes of delivery (online modules, webinars, face-to-face workshop), and duplication with content previously learned. Ninety percent of respondents indicated that they would want to be a Botball mentor again.

7 DISCUSSION

Evaluation findings revealed that there were no major concerns about the module content and format. Responding to the question "What aspects of the mentor training did you find to be least helpful?" forty percent of the participants responded that all of the training was useful:

> *I thought it was all useful. The training and modules had great information.*

A few participants characterized the mentor training overall as "repetitive," suggesting that delivering the same content in the modules, the face-to-face meeting, and the webinars may have been unnecessary for some participants:

> *The mentor training was useful, but often repetitive. More free-form discussion between the mentors would have been helpful.*

In contrast, one participant acknowledged the value in repeating content:

> *I had had experience with mentoring before, but the topics really solidified it and helped it become second nature. Although at times repetitive, it was all useful.*

Taken as a group, the responses did not identify one aspect of the blended learning environment as being more helpful than another. While some participants found it beneficial to interact with other mentors during the webinars, a small number did not. The face-to-face workshop was mentioned by only one participant in the open-ended response; this individual did not think the workshop was important. However, evaluation surveys administered to participants immediately following the face-to-face workshops were overwhelmingly positive.

These preliminary data suggest that the mentors viewed the blended learning environment as a positive experience. Many participants considered the web-based learning through modules or webinars to be effective for them. The mentors were comfortable with the online module format and offered meaningful responses to the embedded open-ended questions. The webinars enabled mentors to form a community of learners, where successful strategies were shared and struggling mentors could ask questions.

There is an extensive literature on learners' experiences in a blended environment (e.g., Bluic et al., 2007). Ausburn (2004) found that adult learners in a blended environment value options, personalization, self direction, variety, and a learning community. In a study of a university accounting course, Cottrell & Robinson (2003) found that students preferred blended learning to the traditional format and classroom time was reduced. Our study used a blended format for adult mentor

training in order to reduce the amount of face-to-face time required. Also, because half of the participants were located on the West Coast of the U.S., and all participants were from different schools, web-based learning was the only practical solution for delivering the content and facilitating ongoing dialogue between the trainers and the participants.

The evaluation to date has uncovered participants' perceptions of the value of different components of the training, but further study is needed to determine possible areas for improvement of the module content and pedagogy. The researchers are planning a focus group to further probe participants' opinions about the value of the modules and to solicit feedback on the design. The feedback will inform revision and refinement of the modules before a public version is made available.

The online module development described in this paper was a part of a larger research study to determine the best approach to mentoring in STEM. The study uses an experimental design to determine if one of three mentoring approaches is more effective in improving STEM-related outcomes for middle school students: mentoring for best practice, mentoring for self-efficacy, or mentoring using a combination of best practices and self-efficacy. The online modules are one component of the intervention being tested, which is fully outlined in Section 5 of this paper. Quantitative analysis of pre/post surveys of mentors and students in the three treatment groups and the control group is in progress and will be reported in a future paper.

8 CONCLUSIONS

The modules described in this paper were designed for a research project focusing on middle school teachers' mentoring skills for student robotics teams. The modules use a research-based approach to train mentors in the theory and practice of group mentoring. The first module is an introduction to group mentoring. Two of the modules focus on best practices in mentoring, including Tuckman's theory of group development. Two additional modules focus on the importance of self-efficacy for student outcomes. They describe the four sources of self-efficacy identified in the research literature and mentoring strategies that support students' self-efficacy. The module content is not specific to grade level or robotics, and instead has broad application for any situation where group mentoring is taking place. This new resource could serve as a valuable tool in other types of academic competitions such as Olympiads.

ACKNOWLEDGEMENTS

The authors wish to acknowledge project team members Profs. Gary Mayer, Susan Thomas, Jerry Weinberg, and Maja Matarić, Mr. Brad White, and Mr. Ross Mead, who provided review comments on early versions of the modules. Ms. Deb Greaney created the interface design in Moodle. Ms. Molly Brooks, Leann Henderson, and Rosemary Morr assisted with coding of open-ended responses. This material is based upon work supported by the U.S. National Science Foundation under Grant No. 1139400. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Ausburn, L. J., 2004. Course design elements most valued by adult learners in a blended online learning environment: An American perspective. *Educational Media International*, 41(4), pp.327-337.

Bandura, A., 1977. Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review*, 84(2), pp.191-215.

Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman. New York.

Bluic, A.-M., Goodyear, P. and Ellis, R. A., 2007. Research focus and Methodological choices in Studies into students' experiences of blended learning in higher education. *Internet and Higher Education*, 10, pp.231-244.

Cottrell, D. M. and Robinson, R. A., 2003. Blended learning in an accounting course. *The Quarterly Review of Distance Education,* 4(3), pp.261-269.

Huang, H., 2002. Toward constructivism for adult learners in online learning environments. *British Journal of Educational Technology,* 33(1), pp.27-37.

Tuckman, B. W., 1965. Developmental sequence in small groups. *Psychological Bulletin*, 63, pp.384-399.

Weinberg, J. B, Pettibone, J. C., Thomas, S. L., Stephen, M. L. and Stein, C., `2007. The impact of robot projects on girls' attitudes toward science and engineering. *Robotics Science and Systems (RSS) Workshop on Research in Robots for Education*, Georgia Institute of Technology, Atlanta, June 30, 2007. Available at http://www.roboteducation.org/rss-2007/,

A Problem-based Learning Case Study for Teaching Voice over Internet Protocol - VoIP

Using Asterisk as a Tool for Teaching VoIP for Information Technology Classes

M. C. Dias[1,2], C. F. Gabi[1], E. P. Rodrigues[1], V. R. Souza[1] and A. Perkusich[2]
[1]*Coordination of Electrical Engineering, Federal Institute of Education, Science and Technology of Paraíba – IFPB, João Pessoa, Brazil*
[2]*Post-Graduate Program in Electrical Engineering – PpgEE – COPELE, Electrical Engineering Department, Federal University of Campina Grande – UFCG, Campina Grande, Brazil*
michel.dias@ifpb.edu.br, cfgabi@hotmail.com, ee.eloisepassos@gmail.com, vicramonee@hotmail.com, percusic@dee.ufcg.edu.br

Keywords: Problem-based Learning - PBL, Asterisk, Voice over IP - VoIP.

Abstract: This paper shows the use of PBL (Problem-Based Learning) technique as a key to learning VoIP in courses like Electrical Engineering and Computer Networks in conjunction with open source and the public domain software called Asterisk which was used to create the scenario of the experiment and the problems presented to the students. In order to make the validation, the experiment was applied to students of Bachelor in Electrical Engineering and in Communication Technology System, undergraduate courses at the Federal Institute of Education, Science and Technology of Paraíba – IFPB, in the Telephony subject, with promising results. The Asterisk software was presented as a useful and flexible tool for constructing scenarios and problems for the teaching of VoIP technologies and the used approach resulted as effective for improving the attainment of the defined learning objectives.

1 INTRODUCTION

In the context of the technological evolution of communications networks, the reality that the telephone networks and data switching networks are converging to an infrastructure, that will allow both voice and data to be transmitted over the same network, is becoming more feasible for the technology and communication professionals on a daily basis. This fact affects professionals from the most diverse backgrounds, who work in the convergent network area, from courses in information technology up to the electrical and telecommunication engineering fields.

The convergence of data networks with telephone networks makes information technology professionals face the challenge of working with scenarios involving IP networks as well as with the existent infrastructure of the traditional telephone networks. Increasingly, the solution to this challenge is related to the use of "Voice over IP" (VoIP) technology in which the phone calls travel through a broadband connection instead of traveling by conventional telephone networks (Keller 2009).

In contrast to such development, the formation of most of these professionals is still based on traditional methodologies in which the teacher is the holder of knowledge and the mass production of labor force is prioritized.

However, methods have been changed and the universities´ great challenge is to provide training courses with the purpose to frame the content according to the student so that he/she can become a technically qualified professional in addition to being able to adapt to frequent changes and demands of the labor market (Silva & Viana, 2013).

With the change of teaching methods, an approach is proposed in which learning is based on practical problems solution (PBL). Such problems are applied in scenarios that simulate a company´s environment and allow students to develop and clarify, in practice, the theoretical concepts that were learned in the classroom, but that were formed in an obscure and abstract way (Fernandes, 2013).

According to Ali and Samaka (2013), problem-based learning is a student-centered, self-directed, inherently collaborative pedagogy where students learn by working in groups through solving

problems and reflecting their experiences. These students are supervised by a tutor or supervisor.

The literature suggests that strategies, in which learning is based on problem solving, are effective in teaching in summarily practical areas, such as information technology. According to Cavalcante and Embiruçu (2013), it is possible to realize how this practice has been established around the world and how it can be applied in engineering courses. The use of VoIP technology has also been established. Furthermore, its importance in the communication systems, that are about to come, is highlighted in Goode (2002). Recently, Dias (et al. 2013) showed that the use of practical experiments can aid in teaching practical concepts of traditional telephony and IP telephony.

This paper will present a case study on the implementation of the Asterisk software as an aid tool related to the practical teaching by using the PBL approach to teach concepts of VoIP technology and its interaction with the traditional telephone system. This methodology was employed in courses in Electrical Engineering and Telecommunication System Technology, but it could easily be applied to any course which had subjects with VoIP technologies in their course programs.

The PBL technique was chosen so as the concept and the motivation could be introduced to the students at the beginning of the subject. This method consists of groups´ choice and of the fulfillment of practical problems which will be solved through the use of software, laboratory practices, theoretical content and the professor´s support. Such method aids the undergraduate students to solve practical questions by themselves (Lamar et. al. 2012).

According to Ribaud and Saliou (2013), PBL can help students to learn with the complexity and perceive that there are no simple responses for problem scenarios, nevertheless learning and life do occur in contexts which can alter the available and possible type of solution.

The problem-based learning method was initially carried out in the Telephony subject with emphasis on the telecommunication area. The students were evaluated in accordance with their theoretical and practical performance, being their critical sense analyzed in relation to the proposed problem.

This paper is organized as follows: section 2 will focus on a bibliographical review on the concepts of VoIP and a presentation of the Asterisk tool; section 3 will describe the materials and methods used in the design of the proposed experiment; in section 4, the results obtained from the experiment will be presented and in section 5, we present the conclusions.

2 VOICE OVER IP AND ASTERISK CONCEPTS

2.1 Voice over IP (VoIP)

Voice over IP is a set of networking protocols that have the function to normalize and regulate the sending of the voice from a source to a destination by using TCP/IP data networks (Keller, 2009). That is, an analog voice signal is converted into a set of digital signals, which is then sent through an internet connection in the form of packets with IP addressing.

The main difference between VoIP and traditional telephony is related to the way the voice is transported. This difference suggests that the only requirement to use VoIP technology is concerned with the use of a TCP / IP connection between two points with end to end delay less than 150 ms. This requirement creates some unique advantages to VoIP, such as (Keller 2009; Goode 2002):

- Cost reduction: expenditure decrease with traditional telecom operators and the use of softphones instead of conventional phones;
- Unique infrastructure: the convergence of voice and data networks will also make the physical network unique;
- Mobility: the branch line must be in a position where you can connect to the Internet;
- The telephone system control: reduces the users´ dependence from the telephone exchange maintenance company; and
- New features: some of them which are not available in the traditional telephony become possible, such as voice encryption during calls.

In VoIP technology, signaling protocols are responsible for determining a standard that specifies the data format and the rules to be followed by data traffic. Moreover, these signaling protocols are used to establish connections, determine the destination and also for issues related to signs such as: ring, caller ID, disconnection, among others. Currently the major signaling protocols for VoIP are (Silva, 2010):

- Session Initiation Protocol (SIP);
- Media Gateway Control Protocol (MGCP);
- Jingle;
- H.248/Megaco; and

- Inter-Asterisk eXchange (IAX).

Taking into account what has been mentioned, it is observed that information technology professionals, who work with communication networks, need practical tools that help them to learn the concepts of IP telephony during their training. The Asterisk, the software already mentioned, however, is presented as an alternative to building practical experiments of similar complexity and low cost for the reality of the job market.

2.2 The Asterisk Platform

The Asterisk software is able to perform the function of a private telephone exchange, which has as one of its primary functions the management of audio transmitted in digital communication channels (Madsen, Meggelen & Bryant, 2011). Asterisk can be used as an extremely powerful and flexible tool designed for the learning of VoIP technologies and protocols once it allows reproducing in laboratories situations and problems only seen in real public or private telephone networks.

Moreover, one of the advantages of working with Asterisk in the classroom is the fact that Asterisk is free software distributed by Digium® that is based on the GPL (General Public License) (Martín, 2009). The free version of Asterisk eliminates the need for a conventional private telephone exchange because its version has no limits of application. Additionally, Asterisk receives users´ contributions from all over the world, making this software always updated (Madsen, Meggelen & Bryant, 2011).

Asterisk's architecture was designed with great care so that there was as much flexibility as possible with regard to the operation of different types of hardware and software (Martín, 2009). Figure 1 shows the Asterisk internal architecture which is formed by a core and specific APIs (Application Programming Interfaces) which support the switching of internal information to PBX.

Information processing in the Asterisk core occurs in such a way that the specific protocols, codecs and hardware interfaces are abstracted from the information. This allows Asterisk to be able to connect to any hardware technology available (either current or future) to perform its essential functions (Silva, 2010).

The functioning and operation of Asterisk are based on the use of modules that the programmer can choose to use or not, depending on the application that he/she is working with. Table 1 describes the main modules for the correct functioning of an Asterisk server (Asterisk, 2010).

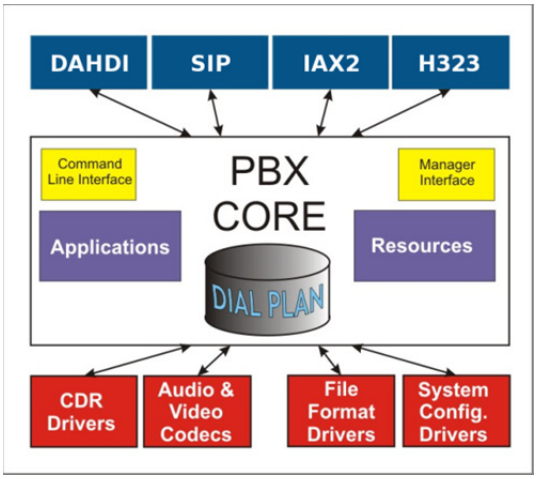


Figure 1: Arquitetura Interna do Asterisk (Asterisk, 2011).

Table 1: Asterisk modules.

Module	Description
Channel Drivers	The channel drivers make the communication with devices outside Asterisk possible by translating the signaling, or protocol, to the core.
Dialplan Applications	This module provides call functionality to the system.
Dialplan Functions	This module is used to set and retrieve parameters of configurations on a call.
Resources	Used to provide resources to Asterisk, like music on hold and call parking.
CODECs	This module is used to encode and decode audio or video so it takes less bandwidth.
File Format Drivers	Used to save media to disk in specific file formats and convert files back to media streams on the network.
Call Detail Record (CDR) Drivers	Used to write call logs to a disk or to a database.
Call Event Log (CEL) Drivers	Similar to CDR, but with details of what happened inside Asterisk during a particular call.
Bridge Drivers	Used by bridging architecture in Asterisk to provide various methods of bridging call media between participants in a call.

When installing Asterisk, the student will be automatically deploying the use of VoIP (Keller,

2009). The software use provides the broadening of learning so that the student can study from the creation of the used VoIP extension lines up to the monitoring of packets sent and the signaling exchange between the terminals. With the purpose of carrying out this investigation, an additional program called Wireshark will be used (Wireshark, 2011).

3 MATERIALS AND METHODS

The experiment was carried out in five phases:

1. Definition of the learning objectives;
2. Construction of a scenario to be set up in a laboratory;
3. Definition of the problems to be presented for the students to solve them;
4. Validation in an undergraduate subject with 60 credit hours, in the minimum, and that had VoIP in its course description; and
5. Evaluation of the obtained results.

The following learning objectives were defined in the first phase:

- Define what VoIP is;
- Differentiate VoIP technologies from the other ones used for establishing telephone calls both in public and private telephony;
- Understand the functioning of SIP, RTP and SRTP protocols; and
- Set up the Asterisk software for establishing telephone calls using VoIP technologies.

In the second phase, the scenario was defined and constructed to be used in the phase related specifically to PBL. The chosen scenario encompasses studies on VoIP with analog terminals and IPs telephones, as shown in Figure 2.

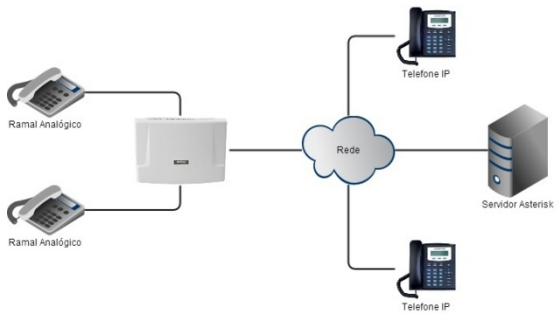


Figure 2: Scenario of the proposed solution.

The objective of this scenario is to provide an environment where students can understand the basic operation of a server set up with Asterisk and that they can, through the Asterisk configuration, make calls between IP terminals that are connected to the server. Subsequently, to understand how the signaling and voice traffic occur on the TCP/IP network, a network monitoring is carried out, in which calls and SIP protocols are captured and RTP/SRTP are analyzed.

When assembling the scenario, the following pieces of equipment available in the IFPB Telephony and Convergent Networks Laboratory were used: Computer with 4GB RAM, processor Intel QUAD CORE, plate FXS / FXO, E1 board, Impacta 68 Intelbras hybrid telephone exchange and Grandstream GXP-1200 VoIP phones.

After defining the scenario, the Linux operating system is installed on a computer which serves as a platform for the use of Asterisk. Figure 3 shows the Asterisk console after it is installed on Linux. It is through the console that the main information regarding Asterisk operation is accessed as well as it is possible to give operation commands to the software.

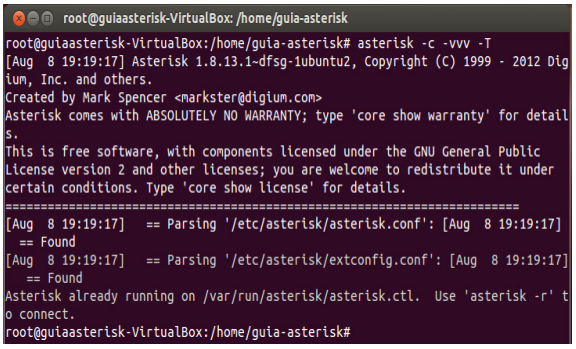


Figure 3: Initiation of Asterisk Server.

Then, it is necessary to create the SIP communication channels so that the Asterisk server can identify each extension. Creating SIP channels with Asterisk occurs by editing the sip.conf configuration file. The code segment below shows how to create four SIP extensions.

```
[general]
Bindport = 5060
Bindaddr = 0.0.0.0
disallow = all
allow = alaw
language = en_US

[commom_to_branches](!)
type = friend
context = branches
host = dynamic
```

```
[2000](commom_to_branches)
secret = 1234
mailbox = 2000

[2001](commom_to_branches)
secret = 1234
mailbox = 2001

[3000](commom_to_branches)
secret = 1234
mailbox = 3000

[3001](commom_to_branches)
secret = 1234
mailbox = 3001
```

In order to have communication among the created extensions, it is necessary to set the Asterisk server dial plan. The dial plan is created by editing the Asterisk extensions.conf. The following code segment shows how the dial plan setup is made for the scenario in Figure 2.

```
[branches]
; Impacta 68 branches

 exten => 2000,1,Dial(SIP/2000,30)
 exten => 2001,1,Dial(SIP/2001,30)

; IP phones branches

 exten => 3000,1,Dial(SIP/3000,30)
 exten => 3001,1,Dial(SIP/3001,30)
```

The use of analog terminals is allowed by setting the central Impacta 68 using specific software, as shown in Figure 4. It is interesting to note that the Impacta 68 central exchange recognizes the Asterisk server as a registration server on the network. Without this setting, it would not be possible to make Impacta 68 extensions to connect with the GXP-1200 IP terminals.



Figure 4: Configuração do Servidor de Registro da Impacta 68.

An IP terminal configuration can be performed on the phone itself or via the web server with the terminal IP address. Figure 5 shows the web server configuration for GXP-1200.

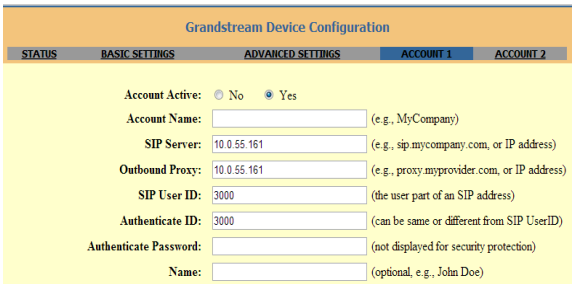


Figure 5: Programming Interface for GXP-1200.

After configuring the extension lines, the terminals send a registration request to the Asterisk server and connect to the server. Figure 6 shows the register of the terminals on the Asterisk server.

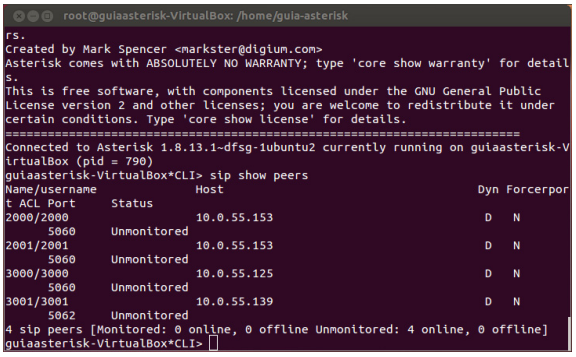


Figure 6: Registration of SIP Channels in Asterisk Server.

In addition to the default implementation, it is possible to activate SRTP protocol on the devices that will provide greater security in the sent packets. Encrypted data, even with packets being captured by network monitoring software, will not give access to the content within them. Therefore, the devices become secure on the network.

In the showed scenario (Figure 2), it is possible to observe how the Asterisk server is able to establish communication with the IP telephony devices and the Impacta 68 Central Telephone through the Internet. Such scenario has great importance to students´ training in the Telephony and Convergent Networks areas, given that its market performance will involve working from assembly to the understanding of the concepts that were applied in practice, in the classroom.

In the next phase, the problems were made up so as to be presented to the students. After researching in workplaces of several telecommunication

companies and interviews with their professionals, the problems were defined to be presented to the students. The main objective was to focus on problems that reflected situations commonly found in professional environments involving VoIP.

For the validation phase, the Telephony subject of the Electrical Engineering and Telecommunication System Technology courses was chosen. Such subject presents all the necessary requirements: class load higher than 60 hours and has VoIP in the course description. In the PBL, the students were evaluated by the professor, by using continuous evaluation in several sessions that the presented problem demanded it, taking into account the learning objectives. At the end, grades were expressed on a scale of 0-100 and they were attributed to each one of the students.

In the evaluation phase, the students answered to a questionnaire of satisfaction and the professor of the mentioned subject carried out a subjective evaluation on the experience. This feedback allowed outlining new actions for improving the used approach. Furthermore, a comparison regarding the repetition and final general mean of the group´s grades in relation the previous semester was accomplished.

4 RESULTS

The mentioned scenario was established several times by students of Telephony discipline, of the Bachelor in Electrical Engineering and the Higher Course in Telecommunication System Technology from the Federal Institute of Education, Science and Technology of Paraíba (IFPB), which covers detailed study of VoIP. This enabled students to carry out the implementation of all practice and successfully absorb the theoretical content previously seen in the classroom.

Using Wireshark to monitor the network, it was possible to capture packets related to the exchange of SIP protocol signaling sent from the terminals to the Asterisk server and vice versa. It was also possible to capture the data stream transmitted among the terminals which uses the RTP protocol. Figure 7 shows an example of the exchange of SIP signaling captured during the experiment.

At the end of the experiment, the students were able to further increase the tab because two IP telephone protocols, RTP and SRTP, were used for security check in sending packets on the network. It was realized that when the SRTP mode is chosen in the terminals, the packets travel on the network in

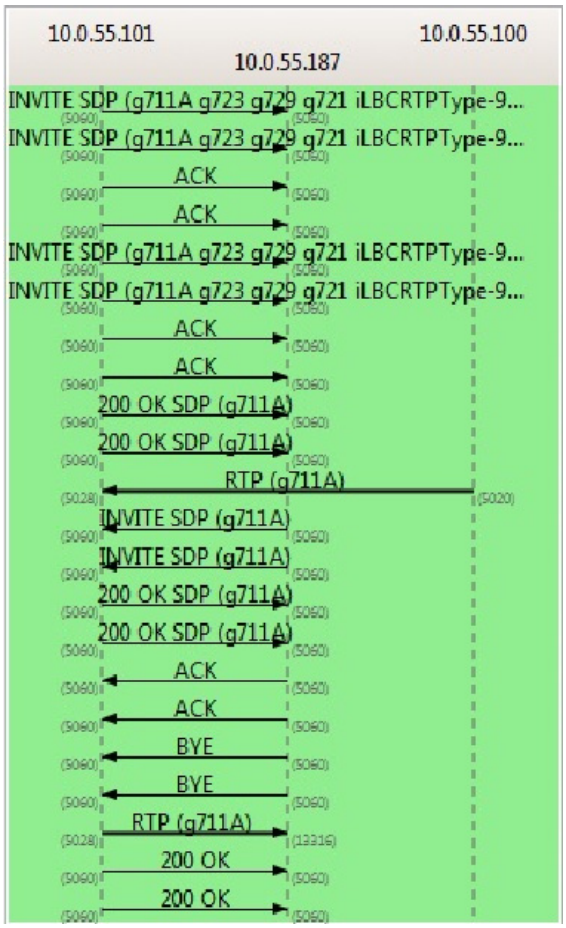


Figure 7: SIP Protocol Signaling Exchanged Between Two Terminals.

the safest way and it will not be possible to hear the content of the calls made, even with the capture of the packets. However for the RTP protocol, it was possible to examine the content of the calls when fulfilling the capture of the packets on the network.

The experiment accomplishment in the classroom allows the understanding of the theoretical content on VoIP technology because the use of Asterisk develops concepts related to both the traditional telephony and IP telephony as well as convergent networks. Furthermore, the use of Asterisk adds to the training of information technology professional a powerful and low cost solution to problems related to voice traffic in computer networks.

After validation, results were encouraging. While comparing with the numbers of the same subject, in the previous semester, the repetition rate was reduced to zero and the final general mean of the group´s grades also increased meaningfully from 61,7 to 81 – an increase of 31,35%. Students were

questioned in relation to the PBL satisfaction level to detriment of the classical approach (in which the professor only transmits knowledge and the student has a secondary role in the learning process) as well as in regard to the general satisfaction with the subject. These results are presented in Figures 8 and 9. In both questionings, the students should mark only one of the following alternatives: very satisfied, satisfied, fairly satisfied, little satisfied or not satisfied.

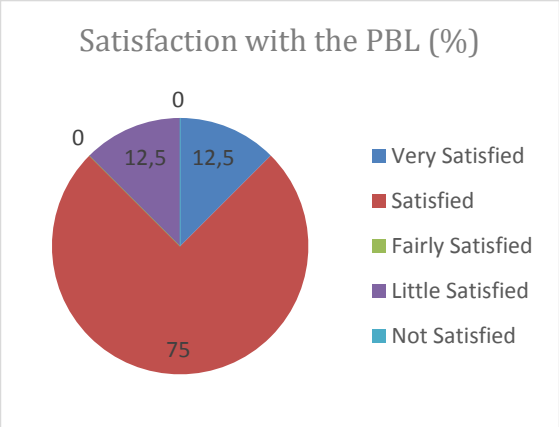


Figure 8: Satisfaction with PBL compared with the classical approach.

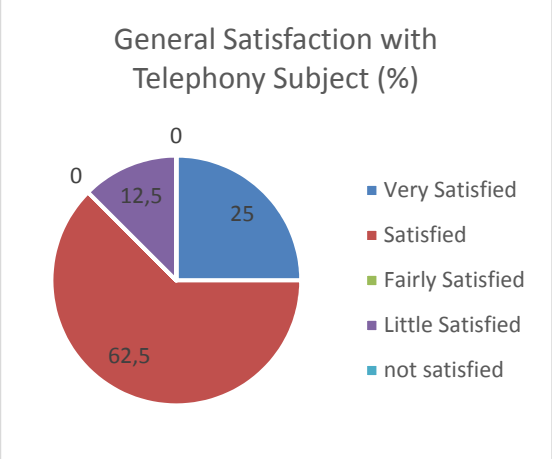


Figure 9: General Satisfaction with the Telephony Subject.

The subjective evaluation of the professor pointed out four important aspects:

1. The proposition of non-trivial problems and with no unique solution increased the challenge level for the students and it acted as stimulus to the participation in the activities of the subject.
2. Students faced difficulty in coping with the diversity of pieces of equipment that make up the scenario of the experiment, though they have had the professor´s explanation about the functioning and operation of such devices at the beginning of the class.
3. There was a noteworthy evolution in regard to the attainment of the general learning objectives in the group from the previous semester.
4. In relation to the PBL, the professor managed to improve the students´ performance, but he/she highlighted the need for teacher´s assistant in case of groups larger than 24 students.

5 CONCLUSIONS

The experience with the PBL using Asterisk as a tool for making up scenarios and problems for VoIP technology teaching was extremely positive both for the professor and the students. In the experimental phase of the mentioned subject, the students´ satisfaction level with the PBL reached 87,5% (satisfied and very satisfied as shown in Figure 8). Furthermore, the evaluation by means of grade and the professor´s subjective perception indicated progress regarding the attainment of the learning objectives on the students´ behalf.

With Asterisk, it was possible to create a reasonable complexity scenario which is present in several telecommunication companies that offer VoIP services; all of this with low cost when compared with the proprietary solutions available in the market. Therefore, providing the students problems that are close to the professional reality, which they will face in the job market, constituted a feasible activity.

As consequence of the analysis of the results, before presenting the problems to be solved for the students´ appreciation, three simple laboratory practices, with the aim to make the students familiar with the equipment and software to be used in the proposed scenario, were introduced. Hence, it is expected to mitigate the problem pointed out by the professor regarding the students´ difficulty in dealing with such devices and the associated software during the resolution phase of the presented problems, in addition to allowing an enhanced focus on the work itself.

Currently, this approach is being used in the same subject and the results will be evaluated in a continuous improvement process. The objective is to consolidate and improve the model so as it can be used for the VoIP technology teaching in the Engineering, Computing and Technology courses.

REFERENCES

Keller, A., 2009. *Asterisk na Prática*. São Paulo: Novatec Editora Ltda.

Zeyad, A. and Samaka, M., 2013. ePBL: Design and Implementation of a Problem Based Learning Environment. In: *Proceedings of the IEEE Global Engineering Education Conference (EDUCON), Berlim, March 2013*. Germany.

Silva, M. and Viana, D., 2013. Variáveis operacionais a serem consideradas no planejamento de disciplinas de projeto para aplicação do PBL nos cursos de engenharia. In: *Proceedings of the 41st COBENGE, Gramado, September 2013*. Gramado: ABENGE.

Fernandes, B. L., 2013. Projetos interdisciplinares: aprendizagem baseada em problemas (PBL). In: *Proceedings of the 41st COBENGE, Gramado, September 2013*. Gramado: ABENGE.

Cavalcante, F. and Embiruçu, M., 2013. Aprendizado com base em problemas: como entusiasmar os alunos a reduzir a evasão nos cursos de graduação em engenharia. In: *Proceedings of the 41st COBENGE, Gramado, September 2013*. Gramado: ABENGE.

Goode, B., 2002. Voice over internet protocol (VoIP). In: *Proceedings of the IEEE, vol. 90, nº. 9, September 2002, pp. 1495-1517 (doi: 10.1109/ JPROC.2002.802005)*.

Dias, M. et al. 2013. Uso do Asterisk como ferramenta de auxílio no ensino prático de telefonia. In: *Proceedings of the 41st COBENGE, Gramado, September 2013*. Gramado: ABENGE.

Lamar, Diego G. et al. 2012. Experiences in the Application of Project-Based Learning in a Switching-Mode Power Supplies Course. *IEEE Transactions on Education,* Vol. 55, no. 1, February 2012.

Ribaud, V. and SALIOU, P., 2013. The cost of Problem-Based Learning: an example in information in systems engineering. In: *Proceedings of the 26th International Conference on Software Engineering Education and Training, San Francisco, May 2013*. San Franciso: IEEE.

Madsen, L., Meggelen, J. V. and Bryant, R., 2011. *Asterisk™: The Definitive Guide*. 3rd ed. Sebastopol: O'Reilly Media, Inc.

Martín, S. G. (2009) *Contribution to Asterisk Open Source Project.* Thesis (MSc). Universitat Oberta de Catalunya.

Finnerty, I. (2009) *Per-extension voip call rating for the Asterisk PBX system.* Thesis (MSc). National University of Ireland.

Silva, V. (2010) *Soluções wireless/voip para redes comunitárias*. Thesis (MSc).Universidade de Aveiro.

Asterisk Project, 2011. *Asterisk Architecture, The Big Picture*. [Online]. Available from: https:// wiki.asterisk.org/wiki/display/AST/Asterisk+Architecture%2C+The+Big+Picture [Accessed 10/10/2013].

Asterisk Project, 2010. *Types of Asterisk Modules*. [Online]. Available from: https://wiki.asterisk.org/ wiki/display/AST/Types+of+Asterisk+Modules [Accessed 10/10/2013].

Wireshark, 2011.*VoIP calls* [Online]. Available from: http://wiki.wireshark.org/VoIP_calls [Accessed 10/10/2013].

Partial Correctness and Continuous Integration in Computer Supported Education

Daniela Fonte[1], Ismael Vilas Boas[1], Nuno Oliveira[1], Daniela da Cruz[1],
Alda Lopes Gançarski[2] and Pedro Rangel Henriques[1]
[1]*Departmento de Informática, Universidade do Minho, Braga, Portugal*
[2]*Institute Telecom, Telecom Sud Paris, Paris, France*
{*danielamoraisfonte, ismael.vb*}*@gmail.com,* {*nunooliveira, danieladacruz, prh*}*@di.uminho.pt,*
alda.gancarski@telecom-sudparis.eu

Keywords: Computer Uses in Education, Problem-based Learning, Competitive Learning, Collaborative Learning, Automatic Grading System, Semantic Program Evaluation, Program Analysis.

Abstract: In this paper we support the idea that students and teachers will benefit from a computer-based system that assesses programming exercises and provide immediate and detailed feedback: students would be able to evolve in the right direction and teachers would follow and assess more fairly their students. This assessment should outperform the typical right/wrong evaluation returned by existing tools, allowing for a flexible partial evaluation. Moreover, we adopt a concept from Agile Development, the *Continuous Integration* (CI), to improve students' effectiveness. The applicability of CI reflects a better monitoring by the teams and their individual members, also providing the ability to improve the speed of the development.
Besides the description of the capabilities that we require from an *Automatic Grading System* (AGS), we discuss iQuimera, an improved AGS that we are working on, that implements our teaching/learning principles.

1 INTRODUCTION

When students start learning Computer Programming, they can experience difficulties to understand the basilar concepts due to a lack of adequate background (Milne and Rowe, 2002) or even adequate proficiency. Actually, Programming is a very complex task specially for beginners who need to have special skills including a high abstraction capability. Programmers need to be capable of understanding the problem statement, analyze it, sketch algorithms and implement them in a programming language. These tasks are equally important and all of them require a lot of effort to be dominated.

In particular, learning a programming language (to code the algorithm and implement the problem resolution) has a lot of complex details that need to be mastered. However, semantic issues involved in any programming language (besides the syntatic pecularities of each one) are the main obstacles to this learning process. We believe that *Problem-based Learning* approach is a good way to diminish this barrier. This is, we advocate that Programming courses should be highly practical; students should solve on the computer practical exercises since the first class, to practice the new language syntax and semantics. Following a problem-based approach, they should learn by practice to deal with all the phases of programming life cycle (since the problem analysis to the program testing). However, to assist all the students in the context of big laboratorial classes (with 20 or 30 students) is almost impossible for the teacher. In real situations, teachers can neither monitor individually all their students nor provide them an exhaustive feedback regarding the exercises they solve. This leads students to give up and fail to reach the course objectives.

From the teacher's perspective, this problem-based approach in Programming courses implies the obligation of manually analyzing and testing the code for each exercise. This time consuming task is neither simple nor mechanical: it is a complex and arduous process, prone to faults, that involves a lot of work. Different teachers may assign different evaluations to the same exercise, due to several factors like fatigue, favoritism or even inconsistency.

Summing up, we recognize that teachers are not able to fully support the students, giving feedback about the mistakes they made in every exercise. This strongly suggests that problem-based learning should be supported by powerful software tools — that is pre-

cisely the focus of the present paper.

Moreover, the changes introduced by *Bologna Process* (Tauch, 2004) have generalized the concept of *Continuous Evaluation* in Education. This technique fosters the work outside classes, providing space for an active and collaborative learning. It also aims at achieving a closer monitoring of each student's progress along a course, by issuing tests and individual or group works. Group-work is another relevant idea: in this way, students are able to share knowledge and take advantage of collective and collaborative activities, like brainstorming or social facilitation; they can also improve soft skills like scheduling, communication, distribution of tasks and leadership (Forsyth, 2009). On the other hand, grouping students in a team provides an efficient solution for teachers to assess multiple students through a singular assessment process, with a significant reduction of the amount of work required compared to individual evaluation (Guerrero et al., 2000).

This new trend also introduces some difficulties both for students (that need to comprehend how to follow it), and for teachers (that must find a just way to assess individually each student). These troubles can, once again, be overcame resorting to the support of intelligent software tools, like Automatic Grading Systems (AGS), as will be discussed along the following sections (this topic is the second motivation for the work here described).

After this motivation and before conclusion, the paper has 3 sections. Section 2 discusses in general the support that computers can bring to education. In Section 3 we introduce our proposal — the main characteristics that an effective AGS should have. Then, in Section 4 we show the architecture of iQuimera, an AGS that we are developing to realize our proposal.

2 COMPUTERS IN EDUCATION

Since a long time ago, computer scientists and software engineers are working on the development of tools to help teachers and students in the teaching/learning process; this gave rise, in the eighties, to the so called area CAE (Computer-aided education) that also involved people from the Education Science and Psychology fields. Two large classes of computer applications were then developed: *tutoring systems*, to help students learning (Natural and Exact Sciences); *learning management systems*, to help teachers to deal with students registry and follow-up.

The evolution of *tutoring systems* went in the direction of the development of tools to assess individually students. In a first phase, using tests (in the form of questionnaires, or similar devices) previously made by teachers; in a second phase, generating automatically those tests selecting questions from a repository or creating them from templates.

Later, with the development of computer communications and the birth of computer networks, people started the construction and use of the so called *e-Learning systems*, on the one hand combining both classes of tools, and on the other hand enabling the distance learning process.

As previously said, in this paper we are interested in a new generation of tools called *Automatic Grading Systems* (AGS) (Matt, 1994; Leal and Silva, 2003; Hukk et al., 2011) devoted to the support of Computer Programming courses. They can be distinguished according to the approach followed to evaluate the submitted program: *static* or *dynamic* analysis. *Dynamic* approaches depend on the output results, after running the submitted program with a set of predefined tests. The final grade depends on the comparison between the expected output and the output actually produced. *Static* approaches take profit from the technology developed for compilers and language-based tools to be able to gather information about the source code without executing it. The improvement of dynamic testing mechanism with static techniques (like metrics or style analysis), led to a new generation of hybrid systems such as CourseMaker, Web-CAT, eGrader, AutoLEP or BOSS that combine the best of both approaches. This symbiosis keeps providing immediate feedback to the users, but enriched by a quality analysis – which is obviously a relevant extra-value.

In the past few years, with the spreading of Agile Methodologies (Beck, 2001), some experiments (Hazzan and Dubinsky, 2003; Jovanovic et al., 2002) were conducted aimed at combining two topics: Agile Development and Education. In this context, the implementation of *Continuos Integration* (CI) (Kim et al., 2008) concept within classrooms seems to be perfectly acceptable as a way of tracking the real evolution of the students in what concerns their programming skills and knowledge about the essential programming concepts, as will be discussed in Section 4.4. With this concept, the evaluation becomes truly continuous, as the name and the original idea assumes. Nevertheless, this matter is not yet deeply explored; to the best of our knowledge there are no AGS which implement this concept.

As we discuss in Section 1, current educational methodologies invest on group work and social collaboration between students to improve their overall effectiveness in regard to learning. In fact, group work is already a reality in education, reducing the time spent by teachers on evaluation and improving

the students' effectiveness on learning. What is missing, however, is to take full advantage of its benefits. We will propose in the next section an improvement of an AGS that supports the adoption of CI practices in education, bringing to teachers valuable resources that allow them to follow and assess the students' evolution with more accuracy.

3 A PROPOSAL FOR AN EFFECTIVE AGS

According to the teaching/learning ideas exposed along Section 1, a software tool to support, on the one hand, the *problem-based learning* and, on the the other hand, the *continuous evaluation* (based on group-work) approaches should comply with the following requirements.

From the teacher's perspective, it should alleviate his tedious and error prone (manual) work of marking and grading the programming exercises submitted by students, allowing him to focus on the students' actual needs.

From the students' perpective, it should make possible to have their works automatically evaluated, finely measured and fairly graded in due time (this means, as fast as possible after each submission).

To do that, the envisaged AGS shall provide detailed reports on each submission. Those reports must say if the answer is totally or partially correct, but shall also include information on the code quality (sofwatre metrics) and on eventual code plagiarism. Student's position when compared with other students and evolution rate are other data that shall be considered in the report.

As discussed in Section 2, at this moment available AGS provide several of these basic features. However, these tools still lack some important features, mainly in what concerns the following: the acceptance of semantically correct answers, no matter its lexico-syntatical appearance; the evaluation of partially correct answers; and the ability to follow students' individual evolution in the context of a continous group work.

All of these considerations led us to build the system that will be introduced in the next section.

4 iQUIMERA

iQuimera is an improved version of our AGS first approach – the Quimera system (Fonte et al., 2012) – developed to support the features announced in Section 3. This system is a web-based application to assess and grade students' programming exercises written in the C programming language, either in *learning* or *programming contest* environments. It offers a complete management system, allowing to set up and manage contests; register students and associate them in groups; and follow up and monitor the different activities involved.

With iQuimera extension, we aim at improving our original approach in two important directions:

- *Flexibilization of the Grading process* – on the one hand, being able to accept as correct answers semantically similar to the expected one, independently of their syntactic differences[1]; on the other hand, allowing the definition of a set of partially correct answers that should also be accepted and graded[2].
- *Collaborative learning* – continuously assessing the group work, providing individual feedback about each group element.

The next subsections introduce iQuimera architecture and its main functionalities. First, we describe the basic features, already implemented in Quimera first version, and then we discuss the improvements (listed above) underdevelopment.

4.1 Quimera Initial Functionalities

Quimera is a system capable of evaluating and automatically ranking programming exercises in *learning* or *programming contest* environments, by combining a very complete static source code analysis with dynamic analysis. Thus, Quimera is able to ensure the grading of the submitted solution based not only on its capability of producing the expected output, but also considering the source code quality and accuracy. Our system guarantees that different programming styles will receive different grades even if they produce the correct output and satisfy all the requirements; this can be an advantage on learning environments, because the student is stimulated to find different correct solutions. Teachers will not waste time searching for problems that have only one solution; the approach adopted allows problems that can be solved in different manners. Moreover, they can settle more interesting problems whose solutions are a set of values, instead of a single value (this constraint, typical of the

[1] The structure of the output file produced by the student program should not be taken into consideration when comparing it to the expected output file.

[2] Using the same semantic similarity approach to compare the output produced by the student program with the partially correct outputs.

classic systems, heavily restrict the kind of statements that can be proposed).

Quimera completes its static analysis with a plagiarism detection tool, in order to prevent fraud among submitted solutions (a common issue in learning environments). To the best of our knowledge, this feature is not provided by the tools referenced above.

Quimera also allies a simple and intuitive user interface with several graphics exhibiting various statistics concerning the assessment process flow. These statistic graphics are useful for the students, as they illustrate their individual and overall evaluation and performance rates; but they are also useful for the teachers, as they provide fast views over each student, groups of students working on the same problem, or over groups working on different problems.

Notice that the feedback returned to the programmer immediately after each submission and the possibility to resubmit new solutions, encourages competitive learning. The student, after receiving the information that his submission is wrong, tends to locate the error, rewrite the program to fix it and submit it again.

4.1.1 Dynamic Analysis

Quimera *Dynamic Analyzer* module follows the traditional strict approach: the submitted (source) code is compiled and then executed with a set of predefined tests. Each test is composed of an input data vector and the correspondent output vector. Running the compiled code (submitted by the student) with one input data vector, the output produced is saved as a vector and is compared with the expected output. The student's submission passes that test if both output vectors are strictly equal (the difference between them is null). In this way, the submission under assessment is correct if and only if all the tests succeed.

After the assessment (running the submitted code with all the test vectors), the user can consult the percentage of tests successfully passed. However, he does not know how many tests there are and only has access to the tests included in the problem statement. This is useful to avoid situations of *trial and error* for *test set guessing*.

4.1.2 Static Analysis

Quimera *Source Code Analyzer* module produces a complete report about the quality of the source code submitted through a set of predefined metrics (currently 56 metrics), grouped by five classes, namely: *Size*, *Consistence*, *Legibility*, *Complexity* and *Originality* (as detailed in (Fonte et al., 2012), here ommited for the sake of space).

The different values computed automatically by this module using classic compiler technologies, when combined with the dynamic evaluation (as discussed in subsection 4.1.1), provide a fine grain assessment that is uttermost relevant for the student learning process. However to be useful in the grading process, the numerical values obtained along the metrics evaluation must be appropriately transformed. For the sake of simplicity and space we do not discuss here those transformations.

4.1.3 Automatic Grading Process

The *Grader* module takes into account the results delivered by the *Dynamic Analyzer* and by the *Source Code Analyzer* and combines them using a grading formula. This grading formula considers six different assessment categories and uses appropriate weights (that can be tuned by the teacher according to the learning objectives). Those assessment categories are the five metric classes introduced in the previous subsection (Size, Consistency, Legibility, Complexity, Cloning) plus one that results from the strict dynamic analysis and that mesures the execution success (the number of tests passed and the execution time). *Cloning* is obviously related with the Originality metric, but more pragmatically measures the percentage of duplicated code in the source code submitted.

At present we are using the following combination of weights: *Execution* (75% + 5%), *Size* (2%), *Consistence* (5%), *Legibility* (2%), *Complexity* (6%) and *Cloning* (5%). In this way, we enhance the dynamic analysis: the actual program ability to produce the correct output corresponds to 80%; the source code quality influences the final grading with just a factor of 20%.

The report issued by Quimera after the Grading phase, contains not only the final grade but all details concerned with the two previous marking phases, as depicted on both sides of Figure 1.

On the left side of Figure 1, we can see a report for an answer that only passes in 22% of the tests. This

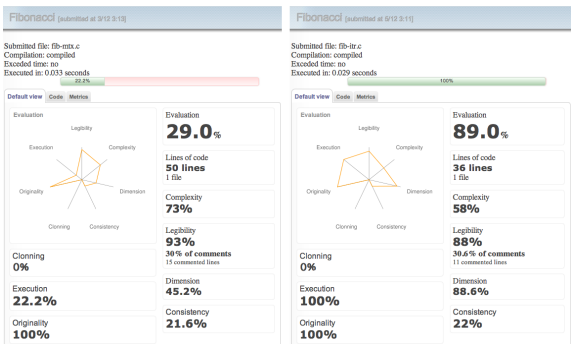


Figure 1: Two assessment report examples.

is a well documented answer (it has 15 commented lines in 50 lines of code) and its *Legibility* has a final score of 93%. The *Complexity* score is owed to the number of used variables and data structures. Since its 50 lines of code exceed the average size of the contest submissions for this problem, this penalizes the final grade on the *Dimension* category (only 45.2%). In the *Consistency* category, its 21.6% are owed to the use of several returns through the code (assuming a maximum of two returns per function as a reasonable limit to the source code consistency) and also to the use of pointers.

The right side of Figure 1 shows a report for a correct answer that passes all the tests. This answer is better documented (11 commented lines in 36 lines of code), but its *Legibility* has a lower final score (88%) once it did not used *defines*, as the other solution. The *Complexity* score of 58% is owed to the implemented *loop*. Its 36 lines of code improve its final grade on the *Dimension* category to 88.6%. In the *Consistency* category, its weak weight of 22% is also owed to the use of several returns through the code.

Finally, it is possible to see that both solutions have not been plagiarized (100% original) and students follow the good practice of not repeating their own code (0% of duplicated code). These assessments led to a final grade of 29% on the first case and a significant 89% on the second case. Quimera also completes this evaluation with radial charts to a quicker and easier comparison of the solution performance in the different grading categories.

4.2 Architecture

iQuimera architecture is organized in four levels, as depicted in Figure 2.

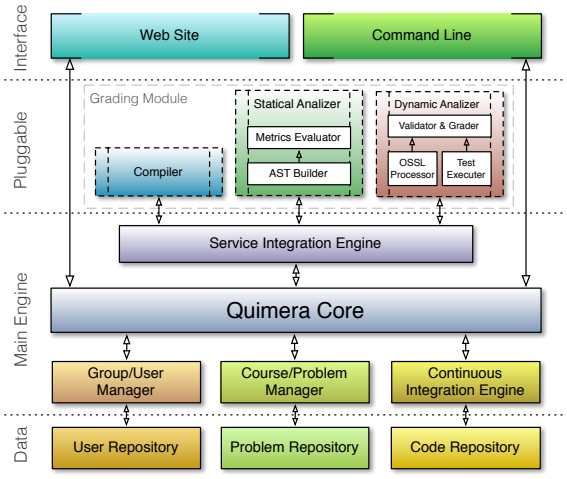


Figure 2: iQuimera Architecture.

- The *Data* level represents the data abstraction layer. It is composed of the modules related with the storage of the data concerned with contests, users, and problems in a relational Database. This level also includes the *Code Repository*, where all the students' submissions are stored.
- The *Main Engine* level works as the system core, linking and controlling the communication between the several components of the system. This level includes: the *Course/Problem* and the *Group/User* managers which are responsible for the interaction with the data associated with the problem and user management tasks; The *Continuous Integration Engine* that is responsible for keeping the *Code Repository* integrity, managing the tasks associated with *Code Repository* update and with the continuous assessment process; and finally, the *Service Integration Engine* that is the API for the communication between the plugins and the system.
- The *Pluggable* level encloses all the plugins currently available in the system. These plugins represent the different tools that can be used in the grading process. They are grouped in three types, namely: *Compilers*, *Static Analyzers* and *Dynamic Analyzers*.
- The *Interface* level is responsible for the interaction between the user and the system. Like the first version, iQuimera is a web-based system with a nice and modern web interface; however we also provide the possibility to active and work with the system via the command line.

4.3 Flexible Dynamic Analysis

As said in Sections 2 and 3, one of the main drawbacks of the existing AGS is their incapability of detecting and correctly grading alternative answers. For instance, they do not support situations where small differences in the order of the output elements do not mean a bad answer (which would be considered correct in a manual assessment).

To overcome this gap we propose a generic model to obtain a more flexible and rigorous grading process, closer to a manual one. More specifically, an extension of the traditional Dynamic Analysis concept, by performing a comparison of the output produced by the program under assessment with the expected one at a semantic level. We aim at allowing not only the specification of which parts of the generated output can differ from the expected one, but also to define how to mark partially correct answers (Fonte et al., 2013).

To implement this model, we perform a semantic comparison between both outputs based on a Domain Specific Language (DSL), specially designed to specify the output structure and semantics – the *Output Semantic-Similarity Language (OSSL)*. The OSSL design also supports the mark of partially correct answers.

For this, we are extending the grading module of Quimera with a *Flexible Dynamic Analyzer* (*FDA*). Figure 3 depicts the architecture of that extension, which is composed of three modules: the *OSSL Processor*, the *Flexible Evaluator* and the *Grader*.

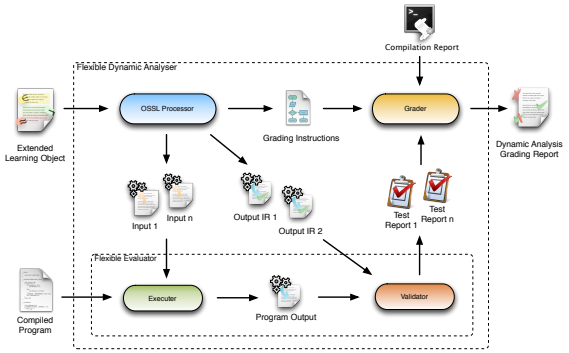


Figure 3: Flexible Dynamic Analyzer Architecture.

The *OSSL Processor* is responsible for producing the set of resources required to execute, validate and grade the submission under assessment. It receives an *Extended Learning Object* containing the problem description, the associated metadata and the *OSSL* specification, and generates the set of *Inputs*, the set of the *Expected Outputs* (through an intermediate representation, IR) and the *Grading Instructions*.

The *Flexible Evaluator* is responsible for executing and validating the submissions. It receives the set of *Inputs*, extracted from the *OSSL* specification by the *OSSL Processor*, and executes the *Compiled Program*. If an execution is successful, the *Flexible Evaluator* module produces a *Program Output* file that is validated against the respective *Output IR* – the intermediate representation of the *OSSL* specification of the expected output, generated by the *OSSL Processor*. This output IR allows to compare (at the semantics level) the meaning of the expected output with the output actually produced. This validation process produces a *Test Report* for each test performed, containing the details about time and memory consumptions and the test results.

The *Grader* module produces a *Grading Report* resulting from the dynamic evaluation performed, concerning the set of *Test Reports* produced by the *Flexible Evaluator* and the *Grading Instructions* provided by the *OSSL Processor*. This *Grading Report* is composed of the details about each individual test report and the submission assessment, which is calculated regarding time and memory consumptions, the weight and score for each test and the number of successful tests. Moreover, if the submission under assessment fails the compilation phase, this grading process is based on the *Compilation Report* provided by the compiler, in order to give feedback about the program under assessment.

Using the *FDA* only requires from the teacher to describe the test set by providing the *OSSL* definition of input and expected output for each problem. This description has only to be written once for each different problem, since this module automatically generates all the data necessary for the student's answers validation.

As an example of usage, consider an exercise where it is asked to write a program that tests which points of a giving set of cartesian coordinates can define a square. As input, the program receives a file with two integers per line, giving the cartesian coordinates of each point. As output, is expected that the student's answer produces the set of four coordinates that defines each square. Considering as input the following set of coordinates: (0,0), (0,10), (10,10), (10,0), (20,0), (20,10), (25,10), (25,0), it is possible to define two squares as depicted in Figure 4.

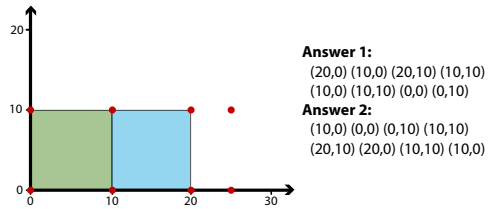


Figure 4: Cartesian representation of the exercise input.

Traditional *AGS* only accept answers containing the groups of four coordinates in a certain order. Also, answers that output only one of the two possible squares are considered incorrect. With the proper *OSSL* definition (as introduced in (Fonte et al., 2013), here ommited for the sake of space), the *FDA* module accepts answers with the correct coordinates for each square in any order or with the squares also listed in any order. Moreover, it can also accept as partially correct answers solutions that only output one of the two possible squares. On the right side of Figure 4, we can see an example of two answers accepted by the *FDA* module, that differ not only on the coordinates order inside each square, but also in the squares order.

We have confidence that the proposed approach is easy to use (*OSSL* allows to specify the output meaning in a simple way) and not difficult to implement.

We also argue that it effectively improves the role of *AGS* as Learning Support Tools, ensuring the interoperability with existent programming exercise evaluation systems that support *Learning Objects* (Hodgins et al., 2002).

4.4 Continuous Integration to Track Students' Evolution

Continuous Integration (CI) is one of the main keys for success within *Extreme Programming*. It grants the quality of the software, by controlling the development process on each small step.

CI takes advantage of the automation of a set of pipelined tasks done in order to improve the productivity in the development (Duvall et al., 2007). It starts by merging the code from multiple team members in an integration machine. This machine stores the latest version of the project and is able to integrate new changes with the existing code, following the synchronous model for team programming (Babich, 1986). This operation may produce collisions in the merging server when two or more files were changed by different team members.

Despite integration may take advantage of usual repository version control tools (like *SVN*[3] or *Git*[4]) to easily handle colisions, CI goes further in the sense that the code is always built and tested upon each commit. A valuable coproduct is the generation of success statistics and reports for every team member who performs an integration (Miller, 2008).

We are enriching Quimera evaluation process by taking advantage of the techniques used in CI, to improve students' effectiveness (Vilas Boas et al., 2013). To the best of our knowledge, CI methodology was never attempted as assessment tool within the educational environment, more specifically, in teaching programming languages to beginners.

Figure 5 depicts the Quimera system extension to include the CI methodology. This process requires the implementation of a *Code Integration Engine* (CIE) with capability to work with some of the common Version Control Systems. New code integration will be automated as much as possible, using the automerge mode of those tools to minimize collisions.

After a successful Integration, CIE builds the code, with the tools already available in the AGS. Along this process, the warnings and errors found are reported. Then, some metrics are evaluated directly from the source code to enrich the final report. As an optimization, files without changes are not reevaluated, speeding up the process.

[3]http://subversion.apache.org

[4]http://git-scm.com

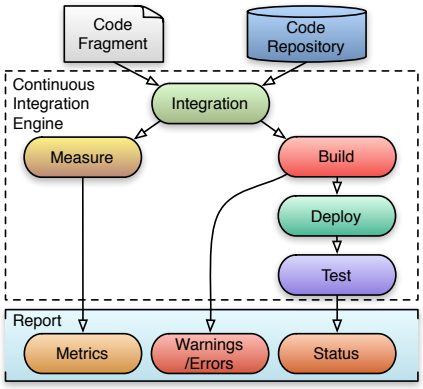


Figure 5: Continuous Integration Engine activity.

Recurring to the AGS, the executable code produced is deployed to a *sandbox* running on protected environment, in order to be tested. This allows to emulate the real environment where the code should be tested and prevents system failure that can come from the execution of malicious code. After deployment, the set of test units is run and a report about the code correctness is generated. This report is the main result of the integration. Finally, another report is generated with details about the status of each step of the process described. This report, available in several formats, is sent to the students and teacher according to their preferences.

iQuimera requires a new set of configurations. To administrators, these new settings are applied at system's initial setup. A configuration routine and documentation will be available to simplify this process. To teachers, the continuous integration process is automatic, so does not require any extra effort. To students, the use of a *CVS* requires machine specific configurations according to their development environment. This may be the most difficult step to the success of its adoption.

As briefly referred in Section 2, there are some good results that arose from the application of *Agile Development* (mainly *Extreme Programming*) methods in educational environments. However, they are applied mostly on advanced years of programming courses due to the programming knowledge level required to follow the complex rules of Extreme Programming. With novices, these will hardly bring positive results, which led us to design this approach to be more suitable for beginners.

We have in mind that this proposal will not soften the steep curve for programming concepts acquaintance, once the basilar concepts will remain hard to grasp. However, with the feedback on continuous evolution on this matter will allow to focus where the problem really exists. Moreover, it also provides a fair

assessment, since students are followed closely and measured for their evolution and implicit effort. All these ingredients contribute to augment the students motivation and enthusiasm.

5 CONCLUSION

We discussed the difficulties inbred in teaching and learning Programming Languages and proposed iQuimera, a system developed in the scope of computer aided education to help making this process more successful through *problem-solving paradigm*. iQuimera extends Quimera, an hybrid (static and dynamic) *Automatic Grading System* (AGS) in what concerns students evaluation approaches.

Our contribution is two fold. On one hand, we propose a more flexible assessment model accepting answers that have syntactic differences (although semantically correct), and also partially correct answers. On the other hand, we propose to apply *Continuous Integration* (CI) principles for a continuous assessment of group elements.

Although iQuimera requires more configuration effort from the three actors enrolled than the traditional *AGS*—more time to specify the problem, test vectors, and grading mechanisms (teacher), to manage all the participants and processes (admin), and to use the system in collaborative mode (students), as well as, more sensibility to tune appropriately the system (teacher)—, it is not an effort necessary each time someone wants to use the system, but it shall be done only once per course or exercise. Thus, we believe that this configuration effort is surpassed by the advantages iQuimera provides with CI and *Flexible Dynamic Analysis* adoption.

We plan (1) to finish the implementation of the *Flexible Dynamic Analyser* and the *CI* engine; (2) to develop new front ends in order for iQuimera to accept other programming languages than C; (3) to extend iQuimera such that it performs with minimal students intervention. Afterwards, we plan to test iQuimera with real users by means of a carefully sketched real case study, in order to verify its usefulness and benefits.

ACKNOWLEDGEMENTS

This work is funded by National Funds through the FCT - Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) within project "Projeto Estratégico - UI 752 - 2011-2012 - Ref. PEst-OE/EEI/UI0752/2011".

REFERENCES

Babich, W. A. (1986). *Software configuration management: coordination for team productivity*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Beck, K. (2001). Manifesto for Agile Software Development. *http://agilemanifesto.org*.

Duvall, P. M., Matyas, S., and Glover, A. (2007). *Continuous Integration: Improving Software Quality and Reducing Risk*. Addison-Wesley Professional.

Fonte, D., Boas, I. V., da Cruz, D., Ganarski, A. L., and Henriques, P. R. (2012). Program analysis and evaluation using quimera. In *ICEIS'12*, pages 209–219.

Fonte, D., Cruz, D. d., Gançarski, A. L., and Henriques, P. R. (2013). A Flexible Dynamic System for Automatic Grading of Programming Exercises. In *SLATE'13*, volume 2, pages 129–144.

Forsyth, D. R. (2009). *Group dynamics*. Wadsworth Publishing Company, Belmont, 5 edition.

Guerrero, L. A., Alarcon, R., Collazos, C., Pino, J. A., and Fuller, D. A. (2000). Evaluating cooperation in group work. In *Groupware CRIWG 2000*, pages 28–35.

Hazzan, O. and Dubinsky (2003). Teaching a software development methodology: the case of extreme programming. In *CSEE&T 2003*, pages 176–184.

Hodgins, W. et al. (2002). Draft Standard for Learning Object Metadata. *IEEE 1484.12.1-2002*, pages i–44.

Hukk, M., Powell, D., and Klein, E. (2011). Infandango: Automated Grading for Student Programming. In *ITiCSE '11*, page 330. ACM.

Jovanovic, V., Murphy, T., and Greca (2002). Use of extreme programming (XP) in teaching introductory programming. *FIE'02*, 2:F1G–23.

Kim, S., Park, S., Yun, J., and Lee, Y. (2008). Automated Continuous Integration of Component-Based Software: An Industrial Experience. In *Automated Software Engineering*, pages 423–426.

Leal, J. P. and Silva, F. (2003). Mooshak: a Web-based multi-site programming contest system. *Software: Practice and Experience*, 33(6):567–581.

Matt, U. v. (1994). Kassandra: the automatic grading system. *SIGCUE Outlook*, 22(1):26–40.

Miller, A. (2008). A Hundred Days of Continuous Integration. In *AGILE '08*, pages 289–293.

Milne, I. and Rowe, G. (2002). Difficulties in learning and teaching programming—views of students and tutors. *Education and Information technologies*, 7(1):55–66.

Tauch, C. (2004). Almost Half-time in the Bologna Process - Where Do We Stand? *European Journal of Education*, 39(3):275–288.

Vilas Boas, I., Oliveira, N., and Rangel Henriques, P. (2013). Agile Development for Education effectiveness improvement. In *SIIE'13*, Viseu, Portugal.

The HumanoidLab
Involving Students in a Research Centre Through an Educational Initiative

Guillem Alenyà[1], José Luis Rivero[2], Aleix Rull[1], Patrick Grosch[1] and Sergi Hernández[1]
[1]*Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Llorens i Artigas 4-6, 08028 Barcelona, Spain*
[2]*Open Source Robotics Foundation, 419 N Shoreline Blvd, Mountain View, CA 94043, U.S.A.*
{*galenya, arull, pgrosch, shernand*}*@iri.upc.edu, jose.luis.rivero.partida@gmail.com*

Keywords: Project based Learning, Mentoring, Humanoid Robots, Introduction to Research.

Abstract: The HumanoidLab is a more than 5 year old activity aimed to use educational robots to approach students to our Research Centre. Different commercial educative humanoid platforms have been used to introduce students to different aspects of robotics using projects and offering guidance and assistance. About 40 students have performed small mechanics, electronics or programming projects that are used to improve the robots by adding features. Robotics competitions are used as a motivation tool. A two weeks course was started that has received 80 undergraduate students, and more than 100 secondary school students in a short version. The experience has been very positive for students and for the institution: some of these students have performed their scholar projects and research in robotics and continue enrolled in the robotics field, and some of them are currently in research groups at IRI.

1 INTRODUCTION

The Institut de Robòtica i Informàtica Industrial (IRI) is a joint research centre between Spanish Research Council (CSIC) and Universitat Politècnica de Catalunya (UPC). In 2007, a SWOT analysis revealed that there was a lack of capacity to involve new students, principally undergraduate and Master students. IRI is located into one of the UPC university campus, but it was relatively unconnected from university activities and the visibility of IRI research areas in the university was very small.

With this in mind, in 2007 the HumanoidLab initiative was created with these objectives:

- Introduce engineering, compute science, and mathematics students to robotics.
- Promote multidisciplinarity in students' education.
- Increase the number of students in the IRI Robotics Laboratory.
- Increase the visibility of IRI research activities in the university community.
- Contribute to the community with open software and hardware projects.

Undergraduate educational robotics can be divided in three areas: robotic clubs, formal university courses, and collaboration into a research project. The HumanoidLab initiative combines a bit all three.

In the first place, some Universities and Research institutions host robotic clubs grouped under different initiatives, like student sections of the IEEE Aerospace and Electronic Systems Society (AESS) [1] or the Board of European Students of Technology (BEST) [2]. These clubs are maintained by students and usually base their activity in organize or participate at national or international robotic competitions, like Eurobot [3] (that takes place from 1998). This competitions propose a set of challenges, and each team designs and builds custom robots with different mechanic, electronic and software solutions. The HumanoidLab wanted to use competitions as motivation tool, but not as primary objective.

In the second place, regular robotic courses involving engineering topics are common (Matari, 2004; Rus, 2006), where simple robots are assembled and programmed with basic reactive algorithms. A good idea to focus the efforts in software development is to use standard and non-expensive robot platforms. In this line, undergraduate experiences using LEGO and AIBO have been reported (Sklar et al., 2007),

[1] http://ieee-aess.org/
[2] http://www.best.eu.org/
[3] http://www.eurobot.org/

and more recently using Roomba based robots like the TurtleBot (Gerkey and Conley, 2011). However, in these platforms is sometimes difficult to embed more sophisticated computer science topics that underlie perception, planning, and control mechanisms in modern robots (Touretzky, 2012). We liked the idea of using robot kits because they provide a quick start platform for people interested principally in programming, but we also wanted to keep the freedom to customize and replace mechanical and electronic parts, and embedded software.

In the third place, there are lots of students collaborating on scientific projects at university departments, but this is usually not performed as a formal educational activity further than master theses and undergraduate works. Besides, departments and research groups have the capacity to participate in competitions that require complex and expensive equipment. Maxwell (Maxwell and Meeden, 2000) reported an intensive summer project where 10 undergraduate students should collaborate to develop a complex robot used to serve hors d'oeuvres in a sophisticated manner. The RoboCup competitions [4], founded in 1997, promote teams involving students not only on the classical soccer competition but more recently the rescue leagues (from 2001) and in the robocup@home (from 2006). On 2010 it engaged more that 3000 participants from 500 different teams. We considered these activities interesting because they are long term projects where hardware and software platforms are evolved during different years and where collaboration and code reuse is promoted. We also wanted that the students could have the opportunity of develop their Master thesis, undergraduate projects, and collaboration grants under the HumanoidLab.

In this paper, the HumanoidLab initiative is described and the results of 5 years of activity are evaluated.

2 METHODOLOGY

The HumanoidLab group decided to use small humanoid robots in all the activities. Time has revealed that this was a key decision because provided us with a distinctive position. In our area of influence, there was a moderate activity in university and bachelor degrees using wheeled robots, but we were the first using humanoids. Humanoids were (and are) a growing topic in the robotics community and give the opportunity to work both in software and hardware and do

[4]http://www.robocup.org/



(a) Original Robonova robot (b) Twiki robot

Figure 1: Twiki robot includes many modifications developed as student projects, as additional d.o.f. at legs, pan-tilt unit, foot pressure sensors, color camera, Gumstix processing unit, improved battery system, and a new C programming library.

it in a wide range of different areas: mechanics, electronics, perception, locomotion, kinematics, manipulation, etc. International market offers the possibility to get a fully working educational small humanoid by a reasonable price (800Eur) and consequently with a reduced funding the HumanoidLab could start their activities. Most of the activities use these educational humanoid robots as a base, and foster their capabilities by adding new features.

The project has 4 project administrators, each one specialized in a different area: mechanics, electronics, low level programming, and perception and planning. Each administrator is the responsible for an activity and for different groups of students. As robotics is multidisciplinar, the tight collaboration between different administrators is usually required and is actively promoted in the group. The HumanoidLab is developing 5 different kinds of activities that are detailed below.

2.1 Small Projects

This is the main activity. It involves a small group of students, between 2 and 3, and it is performed during the scholar University period. We organize meetings between students and some of the administrators to determine the previous knowledge of the students, the topic that they would like to work with, and the time that they could spent. Project administrators propose small projects that fit in the current HumanoidLab interests, and students can also propose their own projects. Once an agreement about the project scope and the execution time has been reached, the project starts with the supervision of one administrator. Each new project provides new capabilities to incrementally obtain more complex and capable robots. In section 3 a detailed list of some of the carried out projects is presented.

2.2 Competitions

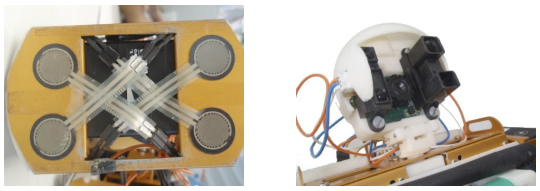
Competitions are not a goal by itself, but an opportunity to put in practice and show the progress of some projects. The HumanoidLab regularly presents some student teams that after the competition write a report (Pons et al., 2010). As it is not a central goal, can happen that the proposed solutions are not the best suited to win the competition, but serve to demonstrate new developments. A clear example is the Twiki robot, who participated at the 2009 CEABOT competition. In the stairs challenge, where stairs have to be recognized and traversed up and down, Twiki used embedded computer vision for obstacle avoidance and pressure sensors to detect downstairs (Pegueroles and Simo-Serra, 2010) (see Fig. 1(b) and 2(a)), while other competitors used simple and effective infra-red sensors. However, it was recognized there as the most technological evolved robot and get the second prize.

In the Spanish national scope, small humanoids competitions are mostly reduced to one or two tournaments per year with events distributed on several days. Each small humanoids contest is composed by several challenges, which can be group into two main categories, depending on what is especially required to the robot:

- Mechanical skills: events designed to probe the robot mechanic design and actuators. An example of this kind its the stairs competition or the sumo contest.
- Perception abilities: events oriented to resolve problems using mainly the robot sensor capacity on the environment. Obstacle race is a good example of this category.

2.3 Introduction Course

UPC University offers the possibility to include a workshop or activity as part of its studies plan. Getting directly into the University studies and participate as an active member has been a good opportunity to have contact with students. A two weeks introductory course/workshop was designed, presented and finally approved by the University. That was the birth of "Introduction to Humanoid Robotics" course. Practice paid a role as important as theory. Participation of students from different schools was one of the strongest proposal points so multidisciplinary teams could be composed. The kind of robots used to perform the practical assignments has evolved from the beginning, and currently from 8 to 10 customized Bioloid robots are used. A maximum of 2 students for each available robot that we have is allowed to join this course at each edition. The course takes 4 hours by day, that is a total amount of 40 hours per course. Past students came from Computer Science, Mechanical Engineering, Electronics, Mathematics and Physics areas.



(a) (b)

Figure 2: Pressure sensors and pan-and-tilt unit on a Robonova robot.

After the great success of this activity we have been asked to propose a shorter introduction activity, of about 2 hours of duration. We have used this shorter activity as outreach in local conferences, secondary school visits, and the European robotics week activities. We have made an effort to develop libraries and tools to facilitate the access to this technology to all kinds of students. Nowadays, the requirements to participate in this course are very basic: knowing the concepts of variable, conditional and loop.

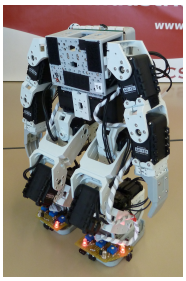
Based on the Introduction course we have also formalized a course in a set of 5 sessions for secondary school teachers, that gives them knowledge and proposes activities to introduce robotics in the secondary school curriculum.

2.4 Curricular Projects

After its participation in some of the HumanoidLab activities, some students decide to perform their curricular projects with us, and sometimes in other robotic groups at the IRI institute. They are basically the final career project (PFC) and the Master Thesis (MTh), but also some undergraduate courses assignments. Some of the works that where performed directly in our group are detailed in Sec. 3.

2.5 Outreach Activities

We have participated in some outreach activities. We have performed some talks at Open Software conferences and to secondary school teachers. we have participated at robot exhibitions like Robot exhibition at Sonar 2010 (Barcelona), IRC 2010 (Korea), and Ficomic (2012). We have appeared on TV and radio programs, and also appeared several times in local newspapers. We regularly participate in the Science Week and in 2012 also in the EU robotics week.



(a) IR foot sensors to detect up and down stairs



(b) Rotating IR sensor at the head

Figure 3: Two different modifications on the Bioloid robot.

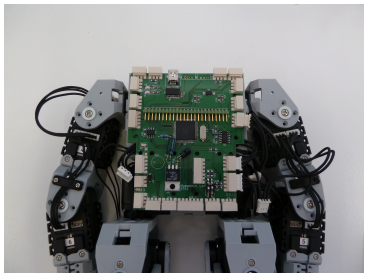


Figure 4: New sensor board for the Bioloid robot.

3 PROPOSED PROJECTS

In five years there have been more that 20 small projects performed. These involve different skills and students have the opportunity to learn different aspects of robotics. A list of some them with code is available on the developers site of the project [5]. Here we list some of the most representative, from the beginning of the project to the latest developments. Along each project it is specified whether it was an students project, a grant, curricular works or coordinators project. We have identified that keeping coordinators motivation is important, and maintain some challenging projects is a good mechanism for such.

3.1 Robonova

The Robonova platform is a 16 dof robot (Fig. 1(a)). It has an embedded micro-controller and can be programmed using a basic-like programming language. It has some analog channels available to use simple sensors.

Inverse Kinematics (students project)**.** This was one of the first activities, defined as a software project. The objective was to define the equations of the Inverse Kinematics (IK) of both arms and legs of the robot, and to implement the corresponding algorithms.

Motivation. IK is important to switch from motions based on pre-computed positions to the complete control over the motion range.

Organization. A group of three students was formed, with complementary knowledge on mathematics an engineering. The project required about 150 hours of work, and 20 hours of supervision based on regular meetings.

Evaluation. Administrators did not correctly evaluated the previous knowledge of the students, and the project required more time than expected. Student were moderately satisfied because they had the opportunity to learn the mathematical background of serial robot kinematics, and the different solutions. One student left the group, and a new project was assigned to the other two in a pure software project following their concerns. An issue was that the project could not be fully completed because the algorithm could not run on the embedded micro-controller. Consequently, the final evaluation was performed in simulation. This motivated the definition of a new electronics project for the addition of a new computing unit.

Extra Dof for Robot Legs (student project)**.** This is a hardware project to modify the leg architecture to allow the inclusion of an extra rotational degree-of-freedom while maintaining almost the same robot height (Fig. 1(b)). This was necessary to not compromise the overall equilibrium.

Motivation. The original Robonova robot has 5 dof at each leg (Fig. 1(a)). Consequently, turns are only possible taking advantage of foot slippage, that is imprecise and with low repeatability.

Organization. The project was assigned to one student with interest in mechanical engineering. The minimum requirements were defined, and the student presented different alternatives that were evaluated, and the best one was implemented on a real robot. The project required about 200 hours of work, and 20 hours of supervision based on regular meetings. The help of the workshop personnel (20 hours) was crucial.

Evaluation. The evaluation was based on the obtained mobility executed by the real robot, and the repeatability of the turning motion. The student found this project as a great opportunity to improve his abilities in mechanical design an related software. He continued in the group, and is now one of the supervisors. The design was published as open source in our website, and several other groups presented similar designs in the next robotics competition.

[5] http://apollo.upc.es/humanoide/

Feet Pressure Sensors (student project)**.** In a previous project a new foot to place force-resistance sensors (FSR) in the bottom and the front, and its associated electronics (Fig 2(a)) was designed. In this project, and thanks to a CSIC grant, a student performed a formal study of the most convenient position and and physical mounting of the sensors, and developed a calibration hardware and software (Barbadillo and Alenyà, 2010).
Motivation. Pressure sensors were used satisfactorily in the skill of climbing up and down stairs. However, we wanted to used them to develop walking algorithms, and more precision was required.
Organization. A single student performed the project. The project was defined taking into account the student background in sensors and electronics. It required 300 hours of dedication. The supervision was based on regular meetings and took 10 hours.
Evaluation. A new evaluation method for us was introduced in this project, based on defining 4 different intermediate checking points with the student. The method worked very well, and let to assess different evolution stages and evaluate the student . It is worth to mention that the knowledge acquired by the student was used in a subsequent research period he performed in another institution (Barbadillo et al., 2011).

3.2 Bioloid

The Bioloid robot is a 18 dof humanoid robot (see Fig. 3) made of modular parts that can be easily changed. Compared to the Robonova, the motors have more torque, and are protected against overload damage. The micro-controller has also basic computation capabilities, and has some analog channels for connecting simple sensors. It come equipped with a small inertial unit intended to aid to the walking motions. The robot is relatively cheap ($\approx$ 800$) so we have 12 of such robots that are used in small projects but also in the Introduction to Humanoids course and related outreaching activities. A project was proposed to add a new extension board (see Fig. 4) to expand the robot controller with different buses and mount new sensors, like a compass. Currently this project has been transferred to a local company that will commercialize it.

Human Motion Transfer (Master thesis)**.** This project uses a depth camera, based on the Time-of-Flight principle, to acquire human motion of an operator and transfer that motions to a Bioloid robot. The student had to implement vision algorithms, 3D segmentation, and implement a communication channel between the vision computer and the robot. Later, with the appearance of Kinect based sensors, the human acquisition part has become more easy, and others have reproduced this work using this sensor (Pfeiffer, 2011).
Motivation. The main objective is to develop a mechanism to create new robot movements that are human-like and easy to create.
Organization. The project was assigned to a student with interest to learn image processing algorithms and interest in control. The project required 400 hours of dedication divided in 3 different stages: 150 hours working in the perception and body segmentation, 50 hours for the adaptation of the IK algorithms and the communication with the robot, 150 hours developing the graphical interface to acquire and reproduce motions, and the rest for writing. It required 35 hours of supervision, and 10 extra hours to solve problems with the image segmentation and robot communication.
Evaluation. This was a Master Thesis project. The student published a report (Siscart, 2011), and the work was publicly presented and evaluated by a jury. The student received the highest qualification. The student was very satisfied because the project entailed closing the perception-decision-control loop, and it had the opportunity to learn perception algorithms.

3.3 Darwin

The Darwin robot (Fig 5) is a 18 dof platform that includes one micro-controller, an embedded PC running Linux, a color camera and an inertial unit. The hardware and the software is open source, and the interaction with the robot is using a C library. Compared to Bioloid, Darwin has a complete PC inside and it is capable of using more complex sensors like color cameras. The architecture is also more robust. However, it costs approximately 10 times the price of a Bioloid. We use this platform for small projects, but it is not possible to have several and use them in the Introduction course. Again, with this platform it is also possible to propose mechanical, electronic, and software projects. For example, in the sensor part a depth camera has been added(Fig 5), and different projects have proposed different solutions to include grippers (Fig 6(a)).

ROS Integration (coordinator project)**.** Currently the ROS framework is widely used and many robots are supported, among others small humanoids. In this project the necessary layers to interface the DarwinOP robot with ROS have been developed.
Motivation. The integration of the robot in the framework enables the use of state-of-the-art algorithms

implementations compatible with ROS, for example for walking (Hornung et al., 2012).

Organization. This project was a coordinator initiative, and assigned to a coordinator because it was a long-term strategical project. It required 200 hours to obtain the first working version. In periodic coordination meetings the evolution was presented and discussed with the other coordinators. The software has been published under LGPL license, following the philosophy of the HumanoidLab. This project is still alive, and requires 5 hours/month of a coordinator for maintenance and to include new features.

Evaluation. The software has been presented in a Darwin competition [6]. Other projects using this software are being proposed, and are used to evaluate and occasionally propose new improvements. The software is also being used by other groups and we receive periodically questions and requirements for improvements that we try to answer.

3.4 Other

Leg Redesign (Master thesis)**.** In this project we investigated how to build a custom pair of legs using powerful motors. The student had to design all the frames and mechanical parts, built the legs, and programmed them to obtain examples of walking gaits. The overall size of the robot was 70cm.

Motivation. Evaluate the complexity and cost of creating custom robot legs to build new robots from scratch.

Organization. The project was assigned to a student with interest in mechanics and hardware abilities. The project required 400 hours of dedication, 30 hours of supervision, and the assistance of the workshop personnel to use the machinery and tools to build the prototype.

Evaluation. This was a Master Thesis project. The student published a report (Cortada, 2010), and the work was publicly presented and evaluated by a jury. He received a high qualification. The evaluation included the demonstration of walking gaits and maximum weight capacity measures.

4 FUNDING

To support the activities of the HumanoidLab we have received different kinds of funding. This has been primarily used to acquire robotic platforms and equipment like computers, additional sensors, and parts, but also to travel to robotic competitions.

[6]http://www.icra2012.org/program/robotChallenge.php

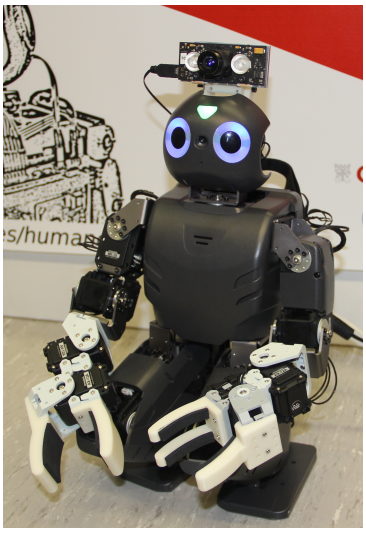
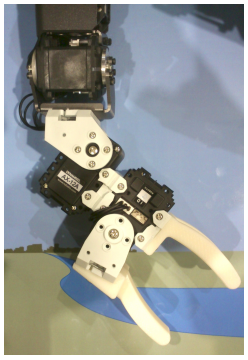


Figure 5: The DarwinOP robot has been modified with grippers and a depth camera. The control system has been completely rewritten to take advantage of the embedded PC/micro-controller architecture.



(a) Simple gripper and additional dof



(b) 3-fingered hand model currently being implemented

Figure 6: Different developments for giving the DarwinOP robot the ability to grasp objects.

Generally students in the lab are volunteers. However, we have obtained funding for grants in three different programs. The first was from the regional government and was intended to foster open source community. We obtained two student grants. The second was from the Spanish Council of Investigation in a program to introduce undergraduate students to research. The third was a program to generate educational contents oriented to secondary school. In total 6 students have obtained a grant.

We started our activities with a reduced number of robots, and we performed some with them small projects and developments. However, this changed when we decided to start the Introduction course and the related outreach activities. More robots were

needed to being able to make these activities, as we planned to use one robot for each two students. We obtained the rest of the robots we needed by buying some new units, but also with a donation from a local retailer.

The HumanoidLab has been mainly funded by a program of the Polytechnic University of Catalonia devoted to strength department activities with 20KEur. Currently, the IRI is funding the developments on the DarwinOP platform of the HumanoidLab with an internal project of 15KEur.

5 EVALUATION

The Robotics Introduction Course has had a great success. 80 students have participated in this intensive 2 weeks activity. After each course the opinions of the students are summarized using reviews. Such reviews reveal that before the course the knowledge about IRI activities was 1 in a scale from 1 (unknown) to 7 (totally known) and before was evaluated as 6. The same applies for the knowledge of the HumanoidLab activities. Concerning the evaluation of the course, students declare that will recommend the course to other colleagues between 6 and 7. Of such students, 5 have performed their PFC in the Humanoid Lab and 1 in another group.

The reviews also ask about knowledge acquired in the course. Students consider that the concepts learned will be useful in their career between 5 and 6. The reviews also include specific questions about the concepts divided in 4 sections: (a) kinematics and actuators, (b) sensors, (c) vision, and (d) robot programming. Questions ask for the knowledge before and after the course, and also the satisfaction after the course. Students declare in (a), (b) and (d) a previous knowledge of 1-2, a present knowledge of 5-6, and a satisfaction of 4-5. Surprisingly, regarding (c) previous knowledge is declared the same 1-2, present knowledge is declared low (3-4), but satisfaction is declared high (5-6). We believe that this is because the robots we are using are quite simple, but computer vision are perceived as very powerful tools. Some students have asked, in a free comment space, for a specific activity about computer vision for robotics.

The Introduction course has been served as a basis for creating a course for secondary school teachers, and also a short version is being currently used for outreach activities of the IRI. More than 100 secondary school students have participated in this activity, always in collaboration with their teachers. Students are in general very satisfied, and their teachers have always expressed that the activity has been very positive and helpful.

More than 40 students have been involved in small projects, and contacted the HumanoidLab thanks to word-of-mouth marketing. This has reported a wide list of improvements to our platforms. Some of these new features have been successfully adopted by other groups working with the same robotic platforms.

Women in engineering and robotics are a minority, at least in Spanish education system. Surprisingly, of the 9 active students nowadays in the HumanoidLab 7 are women. We have been asked sometimes for the reasons by researchers in Sociology, but unfortunately we do not have yet the answer.

We have some special cases that is worth to mention. Three former students of the HumanoidLab have performed its PFC in one research group of the IRI, and currently have become PhD students. One of them is still involved with us and has become administrator. Additionally, 3 students have performed their Master thesis in the HumanoidLab.

We have tracks of 3 former students performing research activities related to robotics in research institutions in other countries.

Currently, we have a DarwinOP robot and 12 Bioloid based robots working. Other robots have become obsolete and no further development is performed on them.

Even if it not an objective, participation in competitions has been very successful. The first year our teams obtained the second and the fourth place at the Spanish humanoid competition. The next year we win the second place again, and in the two last competitions the Humanoid Lab team has obtained the first place. These two teams are actually formed by women.

We believe that, for being useful to the objectives of the HumanoidLab initiative, the participation in competitions should be conceived with the following objectives in mind:

- Multidisciplinary collaboration: autonomous contests requires always from both, hardware and software, development. They represent a good motivation to encourage collaboration between groups of students performing different tasks in order to get a robot working or improve them.
- Promote group spirit: since Humanoid Lab hosts several group of students whom compose different teams, some measures were taken to ensure all people feels like belonging to the same organization. It was the tutor mission to explain new teams that all the previous work they were using was made by their co-workers at IRI, but rivals during the competition. It's a golden rule that any prize won by any of the IRI teams is spent on

improving robots and laboratory material. This avoid any problem or ambition with respect to the money and helps a lot on spreading the work results sharing. For competitions which imply displacement to other cities, the own travel and coexistence helps a lot to create bonds among all students, no matter what team they are.

- Provide long and mid-term explicit goals: the resolution of different competitions events can inspire both, tutors and students, in the search of goals to reach. From an specific competition challenge teams can start working on different projects, from mechanical design to artificial intelligence algorithms, in order to be part of tournament competitors.

6 CONCLUSIONS

The HumanoidLab initiative was funded to introduce students to robotics and give visibility to the research activities of IRI in the student community. After 5 years, nearly 100 undergraduate students have participated in one of the organized activities. Of that, 40 have participated in long term activities. Several of them have continued to work in robotics either in IRI or elsewhere. It is worth to mention that 7 of the current 9 students are women.

Using humanoid robots has revealed to be a good decision, because of the distinction with other activities but also because the inherent problems are challenging. Historically different humanoid robots have been used and improved using open hardware and open source paradigms. The set of performed activities include small projects, curricular projects, competitions, courses and outreach activities.

The synergies between mechanical and electronic engineering, computer science, and mathematicians has been very positive and have been explicitly promoted.

Finally, as it is a volunteer activity it is important to keep coordinator's motivation. One suitable way we have found is to propose challenging projects also to the coordinators that allow to learn from other areas or expand their knowledge in new areas. However, it has to be taken into account that long term projects, suitable to be assigned to a coordinator, can require sustained dedication over the time, like software maintenance.

REFERENCES

Barbadillo, G. and Alenyà, G. (2010). Diseño de un pie para un robot humanoide. Technical Report IRI-TR-10-08, Institut de Robòtica i Informàtica Industrial, CSIC-UPC.

Barbadillo, G., Dautenhahn, K., and Wood, L. (2011). Using fsr sensors to provide tactile skin to the humanoid robot kaspar. UH Computer Science 511, University of Hertfordshire.

Cortada, S. C. (2010). Requirements and design of a humanoid robot for playing soccer. Master's thesis, Universitat Politècnica de Catalunya.

Gerkey, B. and Conley, K. (2011). Robot developer kits. *IEEE Robotics and Automation Magazine*, 18(3):16–16.

Hornung, A., Dornbush, A., Likhachev, M., and Bennewitz, M. (2012). Anytime search-based footstep planning with suboptimality bounds. In *Proc. of the IEEE-RAS International Conference on Humanoid Robots*, Osaka, Japan.

Matari, M. J. (2004). Robotics education for all ages. In *Proceedings, AAAI Spring Symposium on Accessible, Hands-on AI and Robotics Education*.

Maxwell, B. A. and Meeden, L. A. (2000). Integrating robotics research with undergraduate education. *IEEE Intelligent Systems*, 15(6):22–27.

Pegueroles, J. and Simo-Serra, E. (2010). Quadern de treball ceabot 2009 - twiki. Technical Report IRI-TR-10-13, Institut de Robòtica i Informàtica Industrial, CSIC-UPC.

Pfeiffer, S. (2011). Guiado gestual de un robot humanoide mediante un sensor "kinect". Master's thesis, Universitat Politècnica de Catalunya.

Pons, S., Tolos, N., and Troyano, S. (2010). Algorismes i modificacions hardware a la robot dorami per participar al campionat nacional de robòtica humanoide ceabot 2010. Technical Report IRI-TR-10-12, Institut de Robòtica i Informàtica Industrial, CSIC-UPC.

Rus, D. (2006). Teaching robotics everywhere. *IEEE Robotics and Automation Magazine*.

Siscart, M. R. (2011). Algorithms and graphic interface design to control and teach a humanoid robot through human imitation. Master's thesis, Universitat Politècnica de Catalunya.

Sklar, E., Parsons, S., and Azhar, M. Q. (2007). Robotics across the curriculum. In *Proceedings of the AAAI*.

Touretzky, D. S. (2012). Seven big ideas in robotics, and how to teach them. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education*, SIGCSE '12, pages 39–44.

Student Course Evaluation
Class Size, Class Level, Discipline and Gender Bias

Jacob Kogan
Department of Mathematics and Statistics, UMBC, Baltimore, MD 21250, U.S.A.
kogan@umbc.edu

Keywords: World Wide Web, Student Course Evaluation Questionnaire, Instructor Mean, Level Mean, Discipline Mean.

Abstract: Based on approximately 25,000 Student Course Evaluation Questionnaires (SCEQ) covering 16 semesters from Spring 2005 through Fall 2012 and publicly available at the University of Maryland, Baltimore County's website http://www.umbc.edu/oir/sceq/index.html, we analyze the effect of class level and discipline on student responses. We compare the results obtained and conclusions drown with those already reported in the literature.

1 INTRODUCTION

The standard assumption is that Student Course Evaluation Questionnaires are a basic tool for instructor's teaching effectiveness evaluation as well as for institutional improvements of teaching and learning. While more often than not administrators rather confidently believe that the responses to end-of-course assessments represent an accurate description of teacher effectiveness, as a rule, faculty are more skeptical in this regard (Morgan et al., 2003), (Centra, 2003) The following is a summary of the most significant observations concerning teaching evaluations found in the literature:

1. Courses with fewer students receive more positive evaluations than large courses.
2. Humanities courses tend to get better evaluations than science courses.
3. Courses at the advanced level get slightly better evaluations than those at the basic level.
4. Optional courses are better appreciated than mandatory ones.

In a study conducted at the Hong Kong Polytechnic University that focused on the above observations (Kwan, 1999) reached the conclusion that students base their answers on factors external to the course. In a similar line, (Karlsson and Lundberg, 2012) analyzed ninety-eight assessments of faculty from across Swedish universities and concluded that the ratings involve a clear gender and age bias. Younger teachers tend to obtain lower marks in comparison with more senior faculty. Women teachers also consistently receive poorer ratings in comparison with their male counterparts. The effects are worse if the two factors are combined: if you are a young female teacher your evaluations are likely to be significantly below those of a senior male teacher. Gender effect on teaching evaluations is also addressed by (Sprague and Massoni, 2005) with similar conclusions. (Dutceac, 2012) notes: "If a teacher is assigned a mandatory first-year course with one hundred students, she is very likely to get poorer results on the course evaluations than a male colleague teaching a smaller, optional course for the third-year students. And this is regardless of the actual pedagogical skills and competence of the persons in question!"

In this paper we examine some of the above claims by analyzing 24,862 University of Maryland Baltimore County (UMBC) questionnaires generated over 16 semesters. UMBC numerical data supports some of the above observations, and contradicts others. To the best of our knowledge this is the first study covering teaching evaluation data of this magnitude.

2 INSTRUCTOR'S AND UNIT'S RATINGS

Student Course Evaluation Questionnaires (SCEQs) are a basic tool for instructor's teaching effectiveness evaluation. The Student Evaluation of Education Quality (SEEQ) was developed in 1976 by Dr. Herbert Marsh, University of Western Sydney. Marsh

is an internationally recognized expert in the area of psychometrics.

By now various versions of SCEQs have been serving institutions of higher learning around the globe for a long time (Abrami and Cohen, 1990). Course evaluations are incorporated in the process by which Universities seek to improve teaching and learning, and to consider faculty for merit raises, promotion and tenure. The UMBC questionnaire consists of seven sets of items. One set contains general questions that should be applicable to almost all courses. The remaining sets are designed for lectures, discussion, mathematics and science laboratories, seminars, field experience, and self-paced courses. Six questions permit separate evaluation of as many as four instructors. The instructor has the option of administering whichever sets of questions are applicable. This study focuses on general question 9 (G9) "How would you grade the overall teaching effectiveness."

UMBC has been using information collected through SCEQs for a variety of purposes for about forty years. UMBC Student Course Evaluation Questionnaires contain student ratings ranging from 1 (worst) to 5 (best) per each question asked. The SCEQ administration protocol is described in writing, and the Instructor is charged with the responsibility of conducting the survey. The quality of SCEQ scores depends on students' competence to evaluate an instructor and may vary, but a student's evaluation is independent, for example, of the student's GPA. In other terms, ratings assigned by "F" students are as important for instructor's teaching evaluation as those assigned by "A" students.

The ratings per question are averaged out, i.e., the ratings per question are added up and the sum is divided by the number of students who responded to the question (see e.g. (Hardy et al., 1934) where mean evaluations are discussed). This average is named "Instructor Mean."

Along with individual instructor statistics per class/question, SCEQ provides additional statistical indicators, among them "Org Mean" representing a discipline. UMBC computed org means are actually mean averages of the instructor's means. The average scores for a class with one response are weighted equally to a class with numerous responses when "averaging the averages." Instructor Means for classes of different size contribute equally to the Org Mean. Hence the input of large student groups (students in large classes) to the computation of Org Mean is identical to that of small student groups (students in small classes).

To make this point clear we consider a hypothetical "Org" with just two instructors, A and B. While A teaches a small class with one student, B teaches a large class with one hundred students. The single student in the small class rates the "overall teaching effectiveness" (question G9 of the questionnaire, and the main focus of this paper) of the instructor by 5, while each of the one hundred students in the large class responds to the same question with 3. Hence the Instructor A mean is 5, the Instructor B mean is 3, and the Department Mean is $(5+100\times 3)/101 = 3.02$. However the Org Mean reported by on the UMBC SCEQ forms for this hypothetical "Org" would be $(3+5)/2 = 4$.

There are two major deficiencies of the org mean computation currently in place at UMBC:

1. The current method reports instructor B rating 3 significantly below the Org Mean 4, while in fact the Org Mean is only 3.02. The reporting distorts reality.
2. As the example above shows, the opinion of a single student in the small class carries the same weight as that of the 100 students in the large class. In other terms, the voice of one student in the large class worth just 0.01 of the voice of the one student in the small class.

There is no reason for discrimination based solely on the class size.

The results reported in this paper provide means computed in accordance with standard mathematical definition of the arithmetic mean (see e.g. (Hardy et al., 1934), (Hodges and Lehmann, 1964)). The same way means are computed by the University of Maryland College Park (UMCP). Each reference to means computed by UMBC is specifically indicated in the text below.

3 LEVEL MEANS

The UMBC website provides information about 82 different disciplines ("orgs"). We remove 11 "non academic" orgs (with classes such as "Aerobics," "Walking/Jogging," etc.), and focus on students ratings for undergraduate classes (those are 100, 200, 300, and 100 level courses). We compute and graph means for undergraduate courses covering 71 academic disciplines over 16 semesters (Spring 2005-Fall 2012, total of 64 semesters, see Figure 1). The graph shows that the means slowly climb up as levels advance from 100 to 400.

We now partition 71 academic disciplines into 6 clusters:

1. Arts,
2. Engineering & Technology,

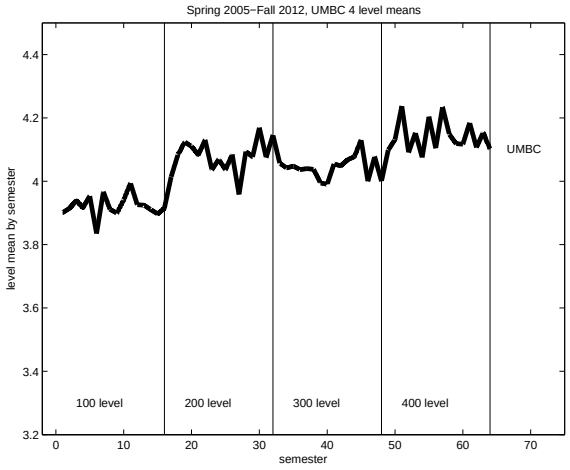


Figure 1: UMBC four level means.

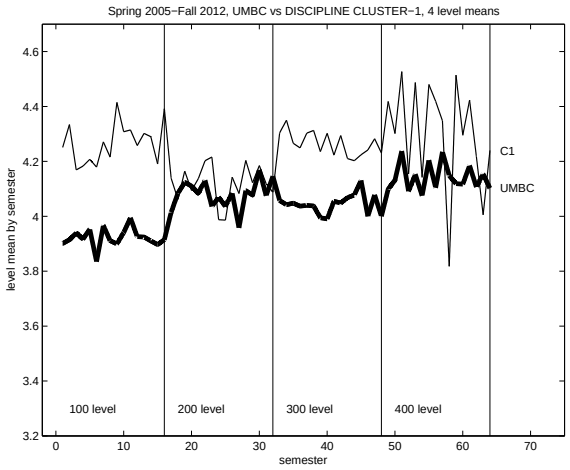


Figure 2: UMBC means vs Arts means.

3. Humanities,
4. Mathematical Sciences,
5. Natural Sciences,
6. Social Sciences.

The means for cluster 1 (Arts) vs. UMBC are shown on Figure 2. The graph reveals that Arts means are higher than the means for the entire university. On the other hand cluster 2 (Engineering & Technology) means, although oscillating around the UMBC means for 100 level classes, clearly fall below the university means as class level grows to 400 (Figure 3). We note that the general trend preserves itself if UMBC means are used as a benchmark, i.e., Humanities and Social Sciences means are above the benchmark, Mathematical Sciences and Natural Sciences are below the benchmark.

The Mathematical Sciences cluster that consists of MATH (mathematics), PHYS (physics), and STAT

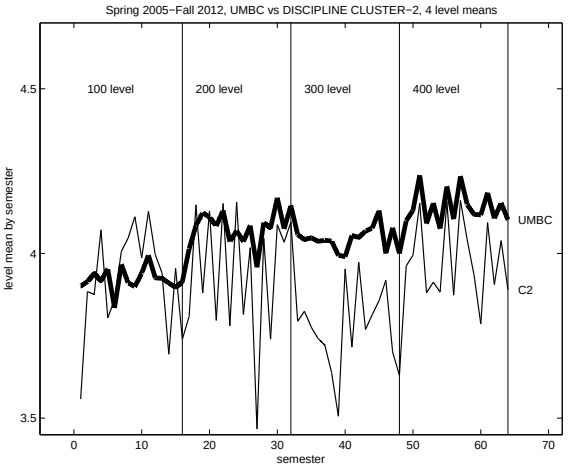


Figure 3: UMBC means vs Engineering & Technology means.

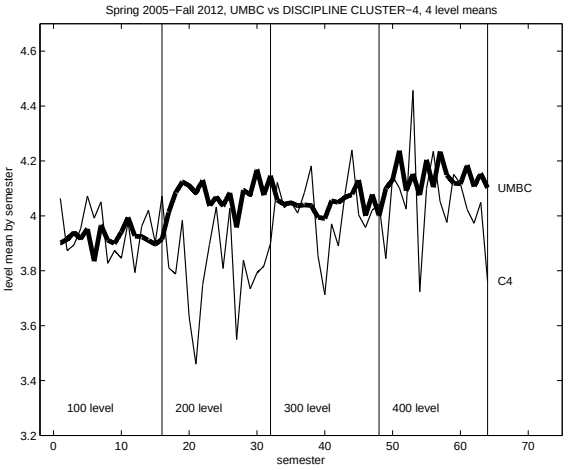


Figure 4: UMBC means vs Mathematical Sciences means.

(statistics) "orgs" exhibits a peculiar behavior (Figure 4). While generally this cluster's means fall under the university means this is not the case for 100 level classes where we observe oscillations around UMBC means, similar to the Engineering & Technology cluster means.

Examination of MATH alone (see Figure 5) shows that MATH 100 level means are higher than the university means. MATH 100 level classes are usually mandatory and of large size, yet student ratings of "the overall teaching effectiveness" of instructors are very high. We note that student ratings for MATH 200, 300, and 400 level classes generally fall under the corresponding UMBC means (Figure 5). Finally we focus on Physics, the "sister" subject of Mathematics. The means of students' ratings Physics vs. Mathematics are shown in Table 6. While at the 100 level classes students rate MATH instruction higher

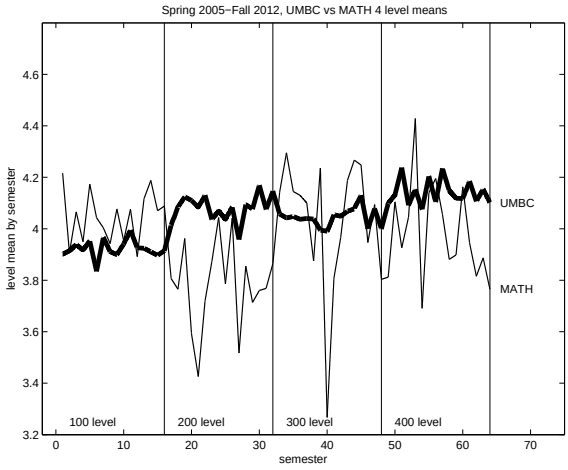


Figure 5: UMBC means vs Mathematics means.

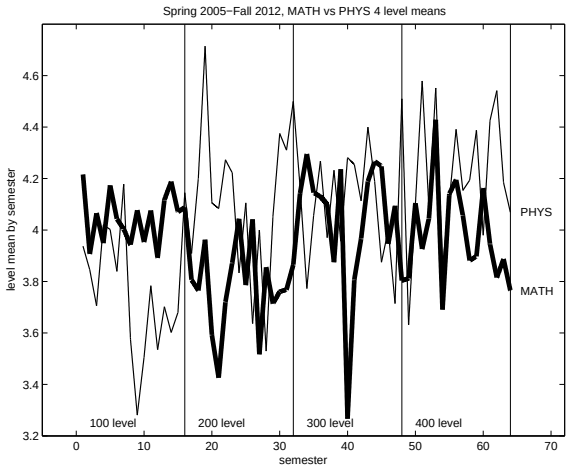


Figure 6: Physics means vs Mathematics means.

than that of Physics, at appears as students mature the ratings abruptly flip at the 200 level. For the rest of students undergraduate life the MATH ratings are generally struggle under the PHYS ratings.

UMBC questioners do not provide data for instructor's gender and rank. In the next section we consider a single "org" for which instructor's gender and rank are available to us.

4 MATH AT UMBC: GENDER AND RANK BIAS

In this section we focus on a single org that manages to generate superior student evaluation results at the 100 level classes. We shall consider one semester only. In Fall of 2012 the Department of Mathematics and Statistics at UMBC was comprised of 31 full time faculty and one post-doc. Approximately $\frac{2}{3}$ of the faculty are Mathematicians, and $\frac{1}{3}$ is made up by Statisticians. Typically (but not always) Mathematicians teach only MATH classes, and Statisticians are involved in STAT instruction only. Five of the faculty are Lecturers (with no mathematical/statistical research responsibilities). For the sake of technical convenience the term "Research Faculty" will denote faculty other than Lecturers (i.e. "Research Faculty" are tenured/tenure–track instructors and the post–doc). The typical teaching work load for Research Faculty is 2 classes per semester, and the teaching workload for Lecturers is 4 classes per semester. There are 7 female and 25 male faculty.

Fall 2012 G9 ("overall teaching effectiveness") rating clearly indicates better evaluations received by female instructors (see Table 1).

Table 1: Mean vs. sex.

sex	mean
male	3.87
female	4.11

At the same time students' ratings of Lecturers and "Research Faculty" show even larger gap in spite of the fact of the heavier teaching load for Lecturers usually conducting instructions in large (100 to 190 students) mandatory 100 level classes. The UMBC Faculty Handbook statement "Effective teaching is absolutely dependent on an active engagement in scholarly efforts" is not supported by the statistics based on SCEQs and provided in Table 2.

Table 2: Lecturers vs. Research Faculty.

rank	mean
Lecturer	4.17
Research Faculty	3.56

5 CONCLUSIONS AND FUTURE STUDY

This paper presents a preliminary analysis of Student Course Evaluations at the University of Maryland Baltimore County. Data provided by the University shows that, contrary to general belief, in some cases student evaluations in large classes are much better than those in small classes, and results of student evaluations of female faculty are better than those of male faculty. The surprise does not stop here. In some of the disciplines teaching evaluation rating for each faculty exceeds the "org" mean reported by the university

(see, for example, Fall 2005, LATN[1]). Clearly there are not that many colleges that can demonstrate a similar result. It should be clear that not everything is rosy. There are still disciplines where the "org" mean reported by the university exceeds all faculty ratings (see, for example, Spring 2012, ARBC[2]).

Future research should investigate gender and faculty rank influence on student ratings. We would like to investigate data pertaining to additional Maryland colleges, and, perhaps, nationwide data. While discipline clustering reported in this study was performed manually in the future studies we intend to apply modern clustering techniques that automatically discover the number of clusters as well as clusters in a given dataset (Kogan, 2007), (Mirkin, 2005).

ACKNOWLEDGEMENTS

The author thanks German Westphal for his input and help with preparation of the manuscript. Insightful conversations with Bradford Sobakewitz are greatly appreciated.

REFERENCES

Abrami, P.C. d'Apollonia, S. and Cohen, P. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology*, 82:2.

Centra, J. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44:5.

Dutceac, A. (2012). Evaluate the evaluation: Course evaluations and external biases. *Inside Higher Ed*.

Hardy, G., J.E., L., and Polya, G. (1934). *Inequalities*. Cambridge University Press, Cambridge.

Hodges, J. J. and Lehmann, E. L. (1964). *Basic Concepts of Probability and Statistics*. Holden-Day, San Francisco.

Karlsson, M. and Lundberg, E. (2012). I betraktarens ögon betydelsen av kön och ålder för studenters läraromdömen. *Högre utbildning*, 2:1.

Kogan, J. (2007). *Introduction to Clustering Large and High–Dimensional Data*. Cambridge University Press, New York.

Kwan, K.-p. (1999). How fair are student ratings in assessing the teaching performance of university teachers? *Assessment & Evaluation in Higher Education*, 24:2.

Mirkin, B. (2005). *Clustering for Data Mining: A Data Recovery Approach*. Chapman & Hall/CRC, Boca Raton.

Morgan, D. A., Sneed, J., and Swinney, L. (2003). Are student evaluations a valid measure of teaching effectiveness: perceptions of accounting faculty members and administrators. *Management Research News*, 26 (7):17–32.

Sprague, J. and Massoni, K. (2005). Student evaluations and gendered expectations: What we cant count can hurt us. *Sex Roles: A Journal of Research*, 53, 11-12:779–793.

[1] http://oir.umbc.edu/files/2013/02/LATN_F05.pdf

[2] http://oir.umbc.edu/files/2013/02/ARBC_S12.pdf

KASS: Korean Automatic Scoring System for Short-answer Questions

Eun-Seo Jang[1], Seung-Shik Kang[1], Eun-Hee Noh[2],
Myung-Hwa Kim[2], Kyung-Hee Sung[2] and Tae-Je Seong[2]
[1]*School of Computer Science, Kookmin University, Seoul, Republic of Korea*
[2]*Korea Institute for Curriculum and Evaluation, Seoul, Republic of Korea*
{akdangz, sskang}@kookmin.ac.kr, {noro, hwa, kelly9147, tjseong}@kice.re.kr

Keywords: Automatic Scoring, Short-answer Questions, Token-based Answer Template, Morphological Analysis, Natural Language Processing, JSON Format.

Abstract: The scoring of short-answer questions in a national-wide achievement test to public school students needs a lot of human efforts and financial expenses. Since we know that natural language processing technology can be applied to replace the manual scoring process by automatic scoring software, many researchers tried to build an automatic scoring system like c-rater and e-rater in English. In this paper, we explored a Korean automatic scoring system for short and free-text responses. NLP techniques like morphological analysis are used to build a token-based scoring template for increasing the coverage of the automatic scoring process. We performed an experiment to measure the efficiency of the automatic scoring system and it covered about 90 to 95% of the student responses with an agreement rate 95% to the manual scoring.

1 INTRODUCTION

Educational achievement test is moving from the offline handwritten platform to an online computer-based assessment (Page, 1966; Valenti, 2003). It causes an automatic scoring method being investigated to save the overall cost of the assessment (Elliot, 2003; Shermis, 2003; Dikli, 2006). Now, automated scoring is one of the interesting research topics in the field of natural language processing in which scoring process includes NLP techniques such as morphological analysis (Kang, 1994). Especially in the case of short-answer questions, lots of different answers to the same question are possible. Therefore, the resources like time, human, finance that is required for assessment can be reduced by the automated system that can handle those questions properly (Sung, 2010; Noh, 2012).

Large-scale exams like CSAT(College Scholastic Ability Test) and NAEA(National Assessment of Educational Achievement) are annual assessment tests in Korea. Automatic scoring for CSAT or NAEA should be fast and accurate. CSAT is very critical to Korean students and the result of the exam determines the entrance of a college or university. Another test TOPIK (Test of Proficiency in Korean) is a Korean language test that is offered four times a year to foreigners.

In this paper, we propose a concept-based scoring method that can evaluate complex questions. Also, we developed a Korean automated scoring system that is suitable for assessing large-scale exams. We performed an experiment on some questions with student answer sets and the result will be discussed in Chapter 4.

2 RELATED WORKS

C-rater is an automatic scoring system for English short-answer questions. The system uses predicate argument structure, pronominal reference, morphological analysis and synonyms to improve scoring accuracy. C-rater system is applied to two studies and results show that about 84% of the responses are assessed identically by the system and human graders (Sukkarieh, 2009; Leacock, 2003). But as mentioned earlier, c-rater is applicable only to English questions.

On the other hand, several methods have been tried for scoring Korean questions automatically. Chung (2009) used cosine similarity between each student response and given model answer. The system assesses responses as correct if critical keywords of the question match (Chung, 2009). Park

(2003) classified open-ended questions to 4 categories; short answer, fill-blank answer, a sentence answer, many sentences answer. The scoring system uses classifying questions, exact matching, applying partial credits, critical keywords and heuristic similarity (Park, 2003).

Kang (2011) classified questions to 6 categories. The scoring system classifies questions then processes them with the proper methods for specific category. Kang's system uses morphological analysis, similarity with heuristics and score calculator (Kang, 2011). Oh (2005) and Cho (2006) used techniques from information retrieval for the scoring system. The system uses semantic kernels, vector space model and latent semantic analysis. Experiments were performed with actual exam papers and the result shows that accuracy of the system was about 80% (Oh, 2005; Cho, 2006).

3 KASS: KOREAN AUTOMATIC SCORING SYSTEM

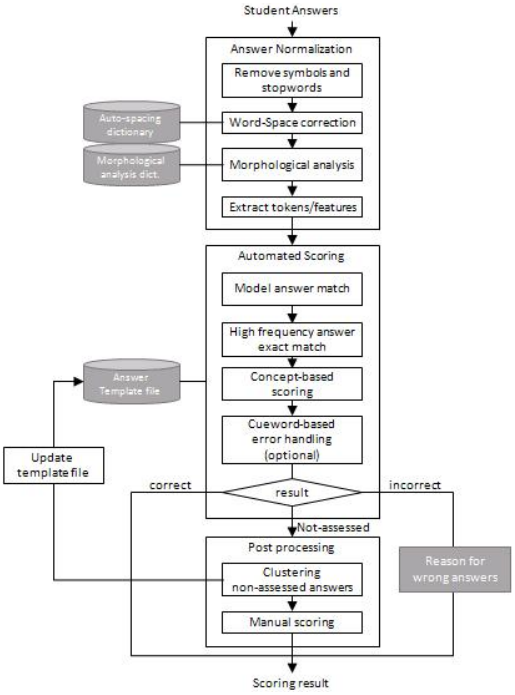


Figure 1: Structure of the proposed system.

The scoring system requires two inputs; student answers to apply automated scoring and model answers with a specific scoring guideline. Model answers are described in an answer template that is constructed by human grader and contains scoring in-formation about model answers, high-frequency student answers, concepts, cuewords and scoring options.

In order to handle automated scoring process, the system should normalize student answers with given options that are defined in an answer template. Then the system proceeds to a scoring step, which also uses an answer template. After a scoring step is finished, the system performs post processing with student answers that are not assessed in the scoring step. Figure 1 shows the structure of the automatic scoring system.

3.1 Answer Templates

Automatic scoring system does not work standalone, but it needs scoring information about a question (Park, 2012). An answer template adopted JSON (JavaScript Object Notation) format which contains detailed scoring guidelines about the question that needs to be assessed. An answer template consists of five parts and the internal structure looks like Figure 2.

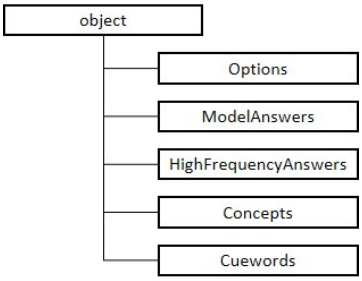


Figure 2: Structure of answer template..

Four options are given for normalize student answers; spelling correction, word-spacing correction, unnecessary symbol remove, and cueword applying option. Options can be changed and stored in the answer template file to be used at the actual scoring stage.

After the options are set, student answers are sorted in the order of frequency and a human gives the score to the high-frequency answers. Scoring information that is generated in this step is also stored in the answer template file. Apart from the "high frequency answers", human graders mark some of the student responses that is to be constructed as initial concepts. Each marked response will be generated automatically as a concept. Finally, human graders can insert some cue words of the question for treating incorrect responses, if the cueword option is set.

3.2 Student Answer Normalization

Although model answers are given for each question, student responses significantly have so many variants of the correct or incorrect answers. So, identifying the characteristics of each question is also important.

Table 1: Example of student responses and result of the normalization.

Answer types	Original text	Normalized text
Model answer	마음을치유 Heal the mind	마음을치유 Heal the mind
response 1	A: 마음을치유 A: heal the mind	마음을치유 Heal the mind
response 2	마음을치유 Healthemind	마음을치유 Heal the mind
response 3	(상처받은)마음을치유 Heal the (hurted) mind	상처받은마음을치유 Heal the hurted mind

For instance, human graders should consider spelling errors about several questions of the Korean language courses. The system should normalize every student responses including model answers with certain considerations. Three of four options are used for normalization stage, except the cueword option. Table 1 is an example of student responses and result of the normalization with all options enabled; spelling error correction, word-spacing error correction and eliminating unnecessary characters or symbols.

3.3 Automated Scoring

Once an answer template file and normalized student responses are prepared, the system is now ready to proceed to the actual scoring phase. Automated scoring process consists of four sequential steps; model answers matching, high frequency answers exact matching, concept-based assessment and cueword-based handling for incorrect responses. A cueword-based task can be skipped by the option. Each scoring step processes every student responses, and then passes the unprocessed responses to the next step.

Model answers and high frequency answers are already evaluated by human graders and normalized by the system. Since the information which are stored in the template file and student responses are already normalized, scoring process can be done with exact matching method.

After exact matching, the system proceeds to the concept-based scoring step. A concept consists of one or more tokens. There are two token types of lexical and grammatical morpheme. A Korean word is split into two tokens through morphological analysis. The system produces concepts from unprocessed student responses then tries to find a concept from the answer template in a conceptualized student response. Concept-based method can handle responses even if part of the answer matches with a concept. Figure 3 shows how the system handles student responses with concepts. The fourth element in the concept has been modified to an asterisk '*' so that any token is matched to this token.

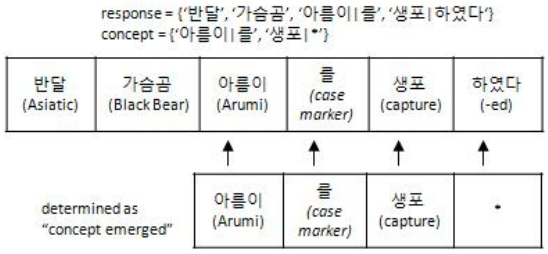


Figure 3: Example of concept matching.

Last step handles incorrect responses with cueword-based methods. Cue word is an essential keyword of the question for writing a correct response. In contrast to the previous step, each response will be treated as incorrect if the response does not include any of the cue words. As mentioned earlier, this step can be skipped by the option. If cueword option is disabled, unprocessed responses after concept-based step will be treated as non-assessed. Non-assessed responses are passed to a post processing module.

3.4 Post Processing

After automated scoring, unprocessed student responses may exist. The purpose of post processing is merging student responses. As a result, a list of merged concepts will be produced. Human graders can examine the produced list and select concepts to create.

Post processing will proceed in the following order; conceptualizing and sorting student responses then merging into concepts. First of all, every student responses are conceptualized. Then post processor sorts student responses by size because merging concept requires same sized responses. If the size of target concepts are same, the system sorts them alphabetically.

Next step is merging the concepts. The system investigates every responses of the same size and picks targets for merging. With the picked responses, the processor performs a comparison,

token by token. If there is an error of less than one or two, the different token is inserted to same position of the other response. Figure 4 shows how concept merging works, where 'Areumi' is a name of a bear. The new concepts that are generated as a result of merging the concepts will be updated in the answer template file.

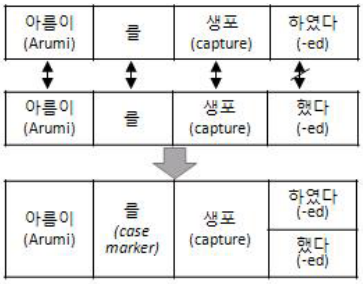


Figure 4: Example of concept merging

4 EXPERIMENTS

We performed experiments to evaluate an efficiency of the scoring system. For the experiments we used about 21,000 student responses from 2012 NAEA. We measured automatic scoring rate per each step. Results are summarized in table 2. The result in table 2 shows that about 80% of the responses are processed in model answers step. 10~15% of the rest is assessed in high frequency answers step, only 5~10% of the responses are passed to concept-based or cueword-based step.

Table 2: Result of the experiments without using cueword option.

ID_Q	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7
# of answers	3018	3013	3017	3017	3012	3023	3035
Model answers	2603	2611	2600	2771	2004	2249	1774
HighFreq answers	296	318	242	121	858	722	1229
Concept-based	66	33	93	54	30	21	23
Non-scored	53	51	82	71	120	31	9

We also compared the results between human grader and the automatic scoring system. The results are in table 3 and the average matching rate was about 95%. The result in table 3 shows that error rates of the system is very low. Note that errors might occur in experiments because the system operator is not an expert of the question.

Table 3: Comparison between KASS and human grader's scoring result.

ID_Q	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7
# of mismatch	4	3	11	4	7	1	12
Matching rate(%)	98.1	98.2	97.1	97.6	95.8	99.3	99.3

5 CONCLUSIONS

We designed and implemented an automatic scoring system for Korean short-answer questions that requires free-text answers. To evaluate the efficiency and accuracy of the system, we performed experiments with 7 questions. The results of experiments show that the system can be applied to a large-scale test like CSAT or NAEA. As a future work, we will focus on some part of the system which is relevant with concepts. Building an answer template semi-automatically will be a part of the work.

ACKNOWLEDGEMENTS

This research has been conducted through R&D program for national assessments of educational achievement funded by Korea Institute for Curriculum and Evaluation (KICE).

REFERENCES

E. B. Page, 1966. The imminence of grading essays by computer, Phi Delta Kappa, pp.238-243.

S. Valenti, F. Neri, and A. Cucchiarelli, 2003. An overview of current research on automated essay grading, *Journal of Information Technology Education*, pp.319-330.

S. Dikli, 2006. An overview of automated scoring of essays, *The journal of technology, learning, and assessment*, pp.1-35.

S. Elliot, 2003. Intellimetric: From here to validity, Automated essay scoring: A cross-disciplinary perspective, pp.71-86.

M. D. Shermis, J. C. Burstein, 2003. Automated essay scoring: A cross-disciplinary perspective, Lawrence Erlbaum Associates, Inc.

C. Leacock and M. Chodorow, 2003. C-rater: automated scoring of short-answer questions, Computers and the Humanities, vol.37, pp.389-405.

J. Sukkarieh and J. Blackmore, 2009. C-rater: automatic content scoring for short constructed response,

Proceedings of the Twenty-Second International FLAIRS Conference, pp.290-295.

T. J. Sung, K. S. Yang, T. H. Kang, and E. Y. Chung, 2010. A survey of computer-based scoring method of free-text or constructed answers in study-advancement evaluation, Korea Institute for Curriculum and Evaluation, Research Report RRE 2010-1.(in Korean).

E. H. Noh, J. H. Sim, M. H. Kim, and J. H. Kim, 2012. A prospect of the developing an automatic scoring system for Korean free-text answers in a large-scale evaluation, Korea Institute for Curriculum and Evaluation, Research Report ORM 2012-92.(in Korean).

E. M. Chung, M. S. Choi, and J. C. Shim, 2009. Design and implementation of automatic marking system for a subjectivity problem of the program, *Journal of Korea Multimedia Society*, vol.12, no.5, pp.767-776.(in Korean).

H. J. Park and W. S. Kang, 2003. Design and implementation of a subjective-type evaluation system using natural language processing technique, *The Journal of Korean association of computer education*, vol.6, no.3, pp.207-216.(in Korean).

W. S. Kang, 2011. Automatic grading system for subjective questions through analyzing question type, *The Journal of the Korea Contents Association*, vol.11, no.2, pp.13-21. (in Korean).

J. S. Oh, W. J. Cho, Y. S. Kim, and J. Y. Lee, 2005. A descriptive question marking system based on semantic kernels, *The Journal of Korean institute of information technology*, vol.3, no.4, pp.95-104. (in Korean).

W. J. Cho, 2006. An Intelligent Marking System based on Semantic Kernel and Korean WordNet, Master Thesis, Hallym University.(in Korean).

I. N. Park, E. H. Noh, J. H. Sim, M. H. Kim, and S. S. Kang, 2012. Answer template description for automatic scoring of Korean free-text or constructed answers, *Proceedings of the 24th Hangul and Korean Language Processing*, pp.138-141.(in Korean).

I. N. Park, E. H. Noh, J. H. Sim, M. H. Kim, and S. S. Kang, 2012. Concept-based automatic scoring system for Korean free-text or constructed answers, *Proceedings of the 24th Hangul and Korean Language Processing*, pp.69-72. (in Korean).

S. S. Kang and Y. T. Kim, 1994. Syllable-based model for the Korean morphology, *Proceedings of the 15th International Conference on Computational Linguistics (COLING-94)*, vol.1, Kyoto, Japan, pp.221-226.

A Blended Health Science Education Graduate Program
Integrating Technology Assignments into Coursework

Sharla King[1,2], Kim Peacock[3] and Lili Liu[4]
[1]*Faculty of Education, University of Alberta, Edmonton, Canada*
[2]*Health Sciences Education and Research Commons, University of Alberta, Edmonton, Canada*
[3]*Centre for Teaching and Learning, University of Alberta, Edmonton, Canada*
[4]*Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada*

Keywords: Blended Learning, Interprofessional Scholarship, Community of Practice.

Abstract: The Master of Education in Health Sciences Education (MEdHSE) program at the University of Alberta, Canada was created to address the need and desire of health professionals and clinical faculty to enhance their teaching, leadership, and educational research skills. A blended learning format was adopted to meet the needs of the learners, all full-time clinical/community educators. A technology integration strategy was developed that included a course focused on technology integration across the curriculum and technology-based assignments for each course. Support for instructors was 'just-in-time' with an educational technology consultant providing individual assistance for each instructor. The educational technologist and graduate students in an educational technology program supported students, with the majority of the assistance occurring early in the program. Students expressed mixed reviews with the technology-based assignments, but overall enjoy the blended learning experience.

1 INTRODUCTION

The increasing complexity of health systems combined with the challenges of managing patients with multiple morbidities, has resulted in a focus on how future health professionals are educated. This includes the preparation of instructors (Frenk et al., 2010). Continuing health professional education refers not only to the enhancement of clinical skills, but also to the enhancement and development of competencies to develop and deliver quality health education. In response to this need, Master's degree programs have emerged worldwide (Tekian and Harris, 2012). A core component that seems to be lacking in most programs is the intentional development of an educational technology strategy that is integrated into courses and assignments.

The Master of Education in Health Sciences Education (MEdHSE) program at the University of Alberta, Canada was created to address the need and desire of health professionals and clinical faculty to enhance their teaching, leadership, and educational research skills. Additionally, to attract working health professionals to the program, flexibility with the delivery of the program was critical. To meet the learners' needs for flexibility and to address the increasing use of educational technology in instruction, a blended learning program was developed that ensured integration of technology throughout the program. This paper will describe the blended program, outline the process for integrating educational technology and describe instructor and student support.

2 BLENDED LEARNING

There is a large and growing body of research that has shown blended learning to be more effective than either fully online or fully face-to-face courses on a number of metrics, including learning outcomes (Means et al., 2010), student satisfaction (Dziuban, Hartman and Moskal, 2004; Wingard, 2004) and levels of student-instructor interaction (Zhao et al., 2005). The University of Central Florida describes blended learning as "a pedagogical approach that combines the effectiveness and socialization opportunities of the classroom with the technologically enhanced active learning possibilities of the online environment, rather than a ratio of delivery modalities" (Dziuban et al., 2004, p 3). While definitions of what constitutes "blended

learning" vary, the courses in the MEdHSE program are closer to the "online learning" end of the blended learning spectrum, to the point where the courses would be considered "online" by some of the more conservative blended learning definitions (Allen et al., 2007).

3 THE PROGRAM

3.1 Development of the Program

The MEdHSE program was developed as a collaboration between the health science faculties and Faculty of Education at the University of Alberta. A program development committee was formed with members' expertise ranging from educational technology, curriculum design, medical education and interprofessional education. The planning process occurred over a two-year period. A decision was made early in the planning process to house the program within the Faculty of Education, rather than one of the health science programs, to ensure the program maintained a strong pedagogical focus across all health science professions. The program extends the knowledge and skills of health science educators in the areas of *educational pedagogy, educational research, and interprofessional leadership* within the collaborative context of communities of practice of professional educators.

3.2 Program Format

The MEdHSE program is designed for health professional educators who work full-time or almost full-time in clinical, community or educational settings and have a role in the education of future health providers or other adult learners. This part-time program is delivered in a blended delivery format consisting of a small number of face-to-face classes and online synchronous and asynchronous sessions. In the initial years of the program, Elluminate (www.elluminate.com) was the online synchronous system used to deliver the course. The University of Alberta recently switched to Adobe Connect (www.adobe.com) for its online or blended programs.

Most of the courses have 1 or 2 face-to-face meetings to supplement the 6 or 7 online classes in that course. The face-to-face time is mainly used for course orientation, lab work, student presentations and some group projects.

One feature of the program is the goal of developing an interprofessional community of practice/learners, which facilitates regular opportunities for engagement, practice and critical reflection. Instructors and supervisors in the program are from both the Faculty of Education and health sciences faculties, with students identifying a health professional with a university appointment to serve as a co-supervisor.

All students enter the program in the course-based stream. At any point during their course of study, students can elect to switch to a thesis-based masters, rather than take the course-based masters. There are eight required courses students must take, regardless of their choice for course-based or thesis-based masters. Each course is delivered in a blended learning format and is 8 weeks in duration. Table 1 provides an overview of the required courses. The one exception is Curriculum Studies in the Health Sciences, which is a 5 day intensive course offered in the spring.

The timing of the face-to-face classes for each course varies. Typically, the first class is a face-to-face class in order to introduce the instructor to the students, with the remaining classes online. For other courses, the face-to-face class comes near the end of the course to facilitate presentations or delivery of student-led content.

Table 1: Summary of required courses.

Year 1
Philosophy of Teaching
Introduction to Methods of Educational Research in the Health Sciences
Learning and Teaching at the Adult Level
Curriculum Studies in the Health Sciences
Year 2
Applied Research Methods
Integrating Technology Across the Curriculum
Assessment and Evaluation in the Health Sciences I
Assessment and Evaluation in the Health Sciences II

3.3 Educational Technology

In addition to a focus on the core courses, a commitment was made to expose students to new technologies used in teaching and learning. This was achieved through two main mechanisms: 1) a course called *Integrating Technology Across the Curriculum* was offered; and 2) an educational technology plan for each course was developed.

3.3.1 Technology Integration Strategy

The philosophy with the technology integration strategy was to build meaningful content together,

both students and instructors. Two goals were identified in the planning of the strategy. First, students in the program would create a series of practical learning objects and/or exemplars for use in their professional practice. These items would be included in the student's portfolio. Second, students in the program would employ a variety of techniques to build a professional learning network (including both their peers in the program, as well as other professionals online). Reflections on and examples of these techniques would be included in the student's portfolio. Although the strategy was developed for the entire program, the adoption was optional for each instructor with 60% of the instructors working with the educational technology consultant to incorporate the strategy into their class.

During the latter stages of the program development phase, the educational technology consultant drafted the technology integration strategy and proposed specific activities and assignments involving technology for each required course. Example strategies from 3 courses, Philosophy of Teaching, Introduction to Methods of Educational Research and Integrating Technology Across the Curriculum, are shared.

3.3.2 Examples of Technology Integration

Philosophy of Teaching: the goal was for students to create a multimedia object that expressed their philosophy of teaching using an online presentation tool, Prezi (www.prezi.com). The students were instructed to develop a personal philosophy of teaching using topics covered in the class. The assignment was to include areas of strength, areas for growth and an action plan for improvement. The student's philosophy was to be connected to the philosophies explored during the course. The purpose of choosing Prezi for this assignment was to allow students to convey these often complex and abstract ideas using a tool that would allow them to represent their ideas in non-linear formats.

Introduction to Methods of Educational Research: the goal in this course was for students to use Twitter *(www.twitter.com) as* a means of keeping up with current research and research methods. At the program's orientation day, students were shown how to create a twitter account and provided with some basic instruction on how to tweet, retweet and search for people or organizations. Additional learning objects about Twitter were created and added to the online course environment to support students who may have needed additional support about how to interact online using Twitter. As part of the course, students were instructed to develop a Professional Learning Network (PLN) with other 'tweeters' in the area of educational research, research methods, health professional education or other relevant field. The students their PLN and used twitter for 5 weeks. At the end of this time, they wrote a reflection on their experiences of using Twitter as a means of developing a professional network and were asked to reflect on the question as to whether this was an effective means of establishing and maintaining a PLN. Several students created a video reflection, rather than a written reflection and one even used an online tool to create avatars to reflect on the use of Twitter with PLNs.

Integrating Technology Across the Curriculum: the goal was to provide the students with a variety of opportunities to work with technology in an applied way, meanwhile utilizing educational theory to support their experiences. For example, before the initial class, students were asked to use Popplet (www.popplet.com), an online tool that allows one to record ideas and concepts in a visual way, to share their ideas about constructivism versus constructionism. Working alone or in groups, students were asked to present to the class on the use of some type of technology. Topics included iPads in the classroom, the use of simulation in teaching and mobile technologies. One optional assignment had the students create an instructional YouTube video.

Student Feedback on the Technology-based assignments: with all of the assignments that used technology, the student reactions were mixed. Some students relished the opportunity to try new mediums to share their assignments, while others wished they could use a more traditional format for the assignments. Others still grumbled at the start of the assignment about the use of technology and then became avid users of the technology. For example, one student ran lunch time tutorials on Prezi for his colleagues at work. Other students actively kept up their Twitter accounts after the course ended and have further enhanced their PLNs. One student in particular was very hesitant to even use the synchronous system at the beginning of the program, however by the 2nd year of the program she was completing optional assignments using YouTube.

3.4 Instructor Development & Support

In order to support the instructors in the blended MEd program, special instructor development sessions were offered. In close collaboration with a educational technology specialist, individual training

sessions with each instructor were offered each term. This was in part due to the diversity of experience and comfort among the instructors with teaching online in a blended format, with the technology tool(s) integrated into their course, and with the scheduling of their courses within the program. The educational technologist consultant working with the program utilized a 'just-in-time' approach for instructor development. When the instructors required technology consulting for course development, or technical assistance in-class, the consultant worked with the instructor on an as-needed basis. Educational technology students also provided additional support for some of the online synchronous sessions.

3.5 Student Development & Support

Student support varied depending on the assignment and instructor. For tools that were used throughout the program (e.g., for student learning portfolios), training time was offered during the student orientation day. In many classes, the educational technology consultant was able to provide an overview or demo of specific tools during class (either online or face-to-face) and then add additional support resources to the online course environment. In the case of more complex technology tools (e.g. Prezi), two or three additional evening training sessions were set up that approximately 25-30% of the students attended. Some assignments used technology tools that required very little instruction. In those cases, online learning objects were created and placed in the online course environment for students to refer to. Regardless of the assignment, students were always welcome to contact the educational technology consultant for just-in-time support.

4 CONCLUSIONS

Health professional educators need to possess competence in not only assessment, teaching and learning, research methods or other traditional core content, but they also must gain skills in using and integrating into curricula a wide range of educational technology.

Preliminary feedback from the students of this MEdHSE program to the instructors, support personal and program director has been positive. Challenges that exist with the technology integration plan include providing off hours support and training to busy professional students, motivating instructors who do not teach these courses more than once a year to adopt the technology assignment, adopting a mix of technology tools that are both innovative and interesting, without being too complex for learners, encouraging professionals who are generally "risk averse" to take risks with trying new technology tools that might positively impact their androgogical practices.

ACKNOWLEDGEMENTS

We wish to thank the instructors in the MEd HSE program for their continued dedication, and the Faculty of Education, in particular the Department of Educational Psychology, for providing unfailing support.

REFERENCES

Allen, I. E., Seaman, J., & Garrett, R. (2007). Blending in: The extent and promise of blended education in the United States. Needham, MA: Sloan Consortium. Retrieved from http://sloanconsortium.org/sites/default/files/Blending_In.pdf.

Dziuban, C., Hartman, J., & Moskal, P. (2004). Blended learning. EDUCAUSE Center for Applied Research Research Bulletin, (2004)7. Retrieved from http://net.educause.edu/ir/library/pdf/ERB0407.pdf.

Frenk, J., Chen, L., Bhutta, Z., Cohen, J., Crisp, N., Evans, T., & Fineberg, H. et al. (2010). Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. *The Lancet, 376,* 1923-1958.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Washington, DC: U.S. Department of Education. Retrieved from http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf.

Tekian, A. & Harris, I. (2012). Preparing health professions education leaders worldwide: A description of masters-level programs. *Medical Teacher*, 34:52-58.

Wingard, R. G. (2004). Classroom teaching changes in web-enhanced courses: A multi-institutional study. EDUCAUSE Quarterly, 27(1), 26-35. Retrieved from http://net.educause.edu/ir/library/pdf/EQM0414.pdf.

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. Teachers College Record, 107(8), 1836-1834.

Using Self-service Business Intelligence for Learning Decision Making with Business Simulation Games

Waranya Poonnawat[1,2] and Peter Lehmann[1]
[1]*Faculty of Information and Communication, Stuttgart Media University, Stuttgart, Germany*
[2]*School of Computing, University of the West of Scotland, Paisley, U.K.*
{poonnawat, lehmann}@hdm-stuttgart.de, waranya.poonnawat@gmail.com

Keywords: Self-service Business Intelligence, Business Simulation Games, Decision Support Systems, Business Intelligence, Management Process, Decision Making.

Abstract: This position paper presents, firstly, the evolution of decision support systems (DSS) and the challenges in teaching in the field of DSS. Secondly, the concepts of management process, decision support technology, self-service business intelligence (SSBI), business simulation games and literature search results on business games associated with DSS are presented. Lastly, we suggest a conceptual framework of using DSS/SSBI on top of business simulation games to support better decision making.

1 INTRODUCTION

Information systems for supporting management decision making, known as *Decision Support Systems*, or DSS, have been evolving since the mid-1960s (Power, 2007). The evolution of DSS concepts remains an important research topic in both industries and universities (e.g., DSS 2.0 Conference, 2014).

Over the last decades, DSS were utilized with some limitations and difficulties, such as heterogeneous data source extraction, multi-dimensional modelling, business analytics, information workers' collaboration, multi-channel user interfaces and massive data visualisation. Meanwhile, the high demand for managing the corporate's information factory (Inmon, Imhoff and Sousa, 2001) brought the modern and powerful DSS concepts and methods onto the DSS stage, which are compromised under the term "*Business Intelligence",* or BI. Starting in the late 90s, Gartner coined the term BI to describe "a set of concepts and methods to improve business decision making by using fact-based support systems" (Power, 2007, p. 11-12). This enables BI applications to function for a wider group of end users and move from the management-focused decision support to the easy-to-use decision support to users at all levels of a company – strategic, tactical and operational. Decision making in companies is necessary for operating and managing highly optimised business processes. Using BI applications with less support from the information technology (IT) departments is called *"Self-Service Business Intelligence"*, or SSBI (Imhoff and White, 2011, p. 5). SSBI is a new BI generation beyond traditional BI technology which needs more IT contribution. Using SSBI tools, users have a variety of personal decision support features and functions, for instance creating, searching, exploring, modelling, analysing, visualising, sharing and collaborating to develop their own ad-hoc BI solutions. The complexity of BI functionalities, therefore, is far more powerful than ever and users are able to use SSBI technology within their desktops or spreadsheet applications with a higher degree of independency from the IT departments.

Nevertheless, the BI-related subjects are typical in the field of Information System (IS) and have been taught for many years (Power, 2007). They are still very popular in the academic world. Subjects such as Information Analytics, Management Information System, Business Intelligence and Business Analytics, are based on DSS/BI concepts. Moreover, Wixom's survey about the BI status in academia (Wixom, Ariyachandra and Mooney, 2013), had a base of 319 professors from 257 universities in 43 countries around the world. There were many BI-related subjects that have been taught in various academic disciplines, for instance Information System, Decision Science, Statistics, Computer Science, Management Information System, Business Analytics, Operations Research,

Supply Chain Management, Economics, Marketing and Accounting.

In the survey's top message, the question about teaching and learning BI was listed as the most challenging issue. Challenges mentioned included: access to data sets, finding a suitable textbook, finding suitable cases and providing realistic experiences (Wixom et al., 2013).

Consequently, the questions about how to teach and learn BI have arised as follows:

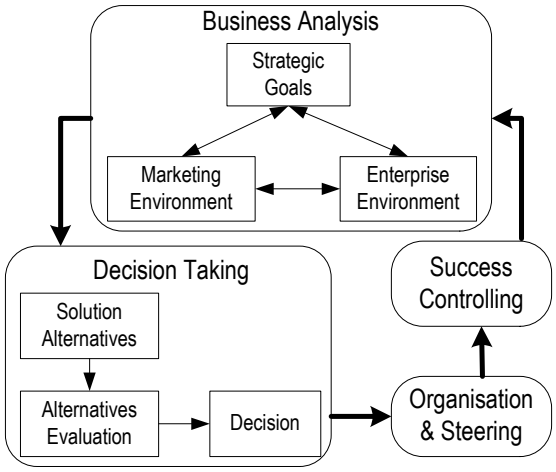
(1) "How can students be taught not only the basic concepts about BI and the handling of a BI tool, but also to select the "right" BI tool for making good decisions in the decision making process?" Since BI/SSBI tools are diverse and often overlap each other, as a consequence they are difficult to use for some users or with a high risk to be overused by other users (Eckerson, 2012).

(2) "What kind of educational platform can be used to teach the effective and efficient usage of BI tools?" Business simulation games are popular and known as one of the most effective education methods, which are widely used for teaching and learning managerial skills, such as making decisions, using management techniques, integrating ideas, applying theory to practice and giving the experiential learning to students (e.g., Ben-Zvi, 2010; Faria, Hutchison, Wellington and Gold, 2009; Lin and Tu, 2012; Wawer, Miloz, Muryjas and Rzemieniak, 2013; Williams, 2011).

In this position paper we suggest a framework to teach and learn BI concepts as a decision supporting method on top of business simulation games. The framework focuses on using SSBI which is a new and powerful BI technology generation for a wide range of decision making by users or information workers. We will focus on SSBI because it gives the opportunities to all kind of users to design DSS/BI models with less IT-technical background needed.

2 MANAGEMENT PROCESS AND DECISION MAKING

DSS/BI technologies are used increasingly to support the management processes, which can be seen as a systematic series of different phases. As an example for management process the following schema will be used to explain the typical management tasks in four phases: Business Analysis, Decision Taking, Organisation & Steering and Success Controlling (Gluchowski, Gabriel and Dittmar, 2008) (see Figure 1).



(Adapted from Gluchowski et al., 2008, p. 21)

Figure 1: Phase diagram for management process.

Phase 1 Business Analysis: the managerial decisions always occur along the business processes, which depend on the context of business objectives, internal and external environment. This initial phase focuses on the (permanent) analysis of situations based on three pillars: *(1) Strategic Goals* – all objectives from all levels of company should be harmonised, not be in conflict and used as a strategic framework to influence and balance with the other two pillars; *(2) Marketing Environment* – such as competition, economic growth and stability and technological advancements, and *(3) Enterprise Environment* – such as availability of resources, organisational culture and structure. All activities, that have an impact or influence on the stability of this system have to be observed, analysed and validated.

Phase 2 Decision Taking: this phase emphasizes the planning for taking decisions and consists of three steps: *(1) Solution Alternatives* – any possible, realistic and relevant alternatives are formulated and collected under a given assumption for any expected future actions; *(2) Alternatives Evaluation* – all collected alternatives have to be evaluated and compared based on the possible risks, feasibility and implications of each alternative, and *(3) Decision* – an alternative is selected out from others which has an acceptable risk and is suitable for a specific business situation for further implementations.

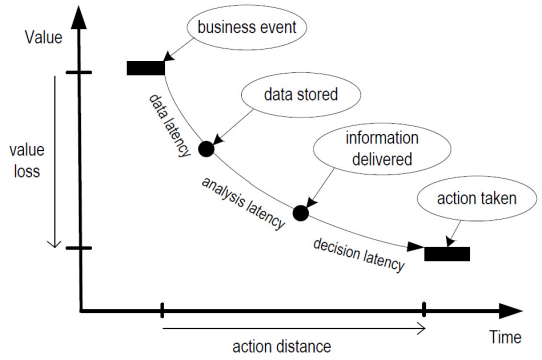
Phase 3 Organisation & Steering: the selected alternative has to be implemented and a course of actions has to be undertaken. Therefore, the organisational structure and project management have to be designed and developed in order to transfer any accountabilities, responsibilities and communication through all hierarchical management

levels during the implementation period.

Phase 4 Success Controlling: the selected alternative is used as the baseline and the actual results are used to measure and compare with the baseline. The variances have to be analysed, which leads to any new actions and restarts the next cycle of the management process.

The business value, which can be gained from management process, depends on the decision making latency or action distance – the distance between the starting point that the business event occurs and the action is taken. The action distance consists of three factors as follows: *(1) data latency* – the time starting from the point that a business event occurs, relevant data are captured, prepared and stored; *(2) analysis latency* – the time for data analysis, information generation and delivery to the proper persons, and *(3) decision latency* – the time to consider and understand all relevant information, make decisions to take the course of action and respond with an intelligent manner (Hackathorn, 2003).

Figure 2 shows the value-time curve – the relationship between the (business) value and time to take the action – which represents as a decay function. The business value decreases rapidly after the business event happens, therefore, the faster to take action, the higher to save business value.



(Adapted from Hackathorn, 2003)

Figure 2: The value-time curve.

3 DECISION SUPPORT TECHNOLOGY

Shim, Warkentin, Courtney, Power, Sharda and Carlsson (2002) stated that computer technology solutions have been used to support complex decision making and problem solving since the late 1950s in terms of DSS and become more significant since the early 1970s. Classical DSS tools have been designed with three main components: (1) the capabilities to access internal and external data, information and knowledge; (2) the functions for modelling and analysing, and (3) the simplified user interfaces to enable interactive queries, reporting and graphing functions. In addition, the research areas of DSS technology typically focus on how to improve the "*efficiency"* of users' decision making and the "*effectiveness*" of decisions.

DSS applications can be used to describe any analytical applications that help managers in planning and optimising business goals and objectives, such as production planning, investment portfolio optimisation, Executive Information System, expert system and Online Analytical Processing (OLAP) (Wixom and Watson, 2010). In addition, data warehouse technology has emerged to handle massive data, operate OLAP and implement dashboard or scorecard applications for DSS (Power, 2007). DSS remain popular in corporate and academic research publications due to the contribution of the four powerful DSS technologies: data warehouse, OLAP, data mining and World Wide Web (WWW) (Shim et al., 2002).

Since the early 1990s, Gartner coined the term BI and the term BI also has been used to describe the analytical and decision support applications. Wixom et al. (2010, p. 14) also defined BI as "a broad category of technology, applications and processes for gathering, storing, accessing and analysing data to help its users make better decisions". The authors also stated that BI plays a critical role, impacts to organisational success, is required to compete in the marketplace and changes from being used by a few specialists to many workers.

In today's economic environment, BI solutions become more important and essential for managing the company intelligently. However, many decisions still are not based on BI because of the limitations to access information and to use suitable BI tools for business analytics. A new development of BI technology called Self-Service BI, or SSBI, offers an environment to support and empower users to create their own ad-hoc BI solutions and making decision faster. The development of SSBI technology is highly growing and the new SSBI functionalities will be launched more into the marketplace (Evelson, 2012; Howson, 2013).

4 SELF-SERVICE BUSINESS INTELLIGENCE

The concept of personal decision support systems is the oldest form of DSS (Arnott, 2008), and the concept of SSBI has been attempted to integrate in BI systems for many years (Mundy, 2013). Originally, the objective of both concepts is supporting personal decision making. In recent years, however, the development of SSBI emerged as a new advanced BI technology in the marketplace in order to fulfil this objective. Some significant drivers for SSBI requirement are as follows: the business needs change constantly and rapidly, the IT departments are unable to satisfy the business users' requirements in timely manner, the slow access to information provided by the IT departments, the business users need to do more analytics and the limitation of IT budget (e.g., Eckerson, 2012; Imhoff and White, 2011; Kulkarni, 2012).

SSBI is defined as "the facilities within the BI environment that enable BI users to become more self-reliant and less dependent on the IT organisation" (Imhoff and White, 2011, p. 5). These facilities focus on four main objectives: (1) to make BI results easy to consume and enhance; (2) to make BI tools easy to use; (3) to make data warehouse solutions fast to deploy and easy to manage, and (4) to make data sources easy to access (Imhoff and White, 2011).

Since SSBI tools are diverse, an appropriate self-service environment can be provided by knowing the types of information workers, the skill levels of different information workers and the tools or fuctions of SSBI they need (Imhoff and White, 2011). Moreover, Imhoff and White (2011) found that business users' skills and the lack of business users' training are two of the top five inhibitors for SSBI.

5 BUSINESS SIMULATION GAMES

Business simulation game is a subset of simulation games which focuses on business content, whereas, the broader definition of simulation game underlying of two concepts: simulation and game. The term *"simulation"* generally refers to "a representation of a real system, an abstract system, an environment or a process that is electronically generated" (Hainey, 2010, p. 44). The term *"game"*, is defined by Hay as "an artificially constructed, competitive activity with a specific goal, a set of rules and constraints that is located in a specific context" (as cited in Wilson, Bedwell, Lazzara, Salas, Burke, Estock, Orvis and Conkey, 2009, p. 2). Cruickshank stated that the term simulation game is used as "one in which participants are provided with simulated environment in which to play" (as cited in Connolly and Stansfield, 2006, p. 466).

Faria et al. (2009) stated that business simulation games have been developed and used as the vehicles for teaching the business concepts for more than 40 years in universities and companies. The major reasons of using business simulation games were as follows: gained experience, strategy aspects, decision-making, learning outcomes and teamwork experience. The advancement of IT provided more opportunities to improve the learning experience and the way to use business simulation games and also to develop a more complex environment. In addition, business simulation games have moved from being a supplemental tool to a central tool and have become a major form of pedagogy for business education.

Several studies stated that business simulation games enable students to learn how to make decisions, manage the business process in a modern enterprise, link between abstract concepts and real world problems and improve quantitative skills (e.g., Ben-Zvi, 2010; Lin and Tu, 2012; Wawer, 2013; Williams, 2011). Furthermore, the new concept of business simulation games, which combines with case-based approaches and experience-based learning theories, results in business simulation games being one of the popular and effective way of education methods (Wawer, 2013).

6 LITERATURE SEARCH

The literature search has been performed to find the empirical studies about business games associated with DSS. The literature search has been done using several online databases – Google Scholar, ScienceDirect, EBSCO, IEEE, Springer, Wiley Online, ACM and Emerald. The terms used for searching from abstracts, titles and keywords, as follows:

("serious games" OR "business games" OR "games-based learning") AND ("decision support system" OR "management information system")

The initial search returned 1,362 results, of which ten articles met the criteria - business games associated with using DSS for making decision – and two added articles were found in the references.

The studies showed that some business simulation games provided decision support tools inside the games and some others used the external decision support tools. The reporting in the business simulation games was often based on pre-defined queries with little flexibilities in using ad-hoc queries. Analytical modules for prediction were restricted to the database of the games. The flexibility for tactical queries and automated decisions were not foreseen. Moreover, business simulation games in the studies were not designed with regard to teaching and learning BI concepts. However, the strength was on teaching business scenario.

7 PROPOSED SOLUTION

This position paper suggests a conceptual framework of using DSS on top of business simulation games to teach and learn decision making (see Figure 3).

In the research project, business simulation games will be used as an educational platform to simulate the business scenario. During the business processes, the DSS tools – SSBI – will be applied for each business activity to support the decision making process.

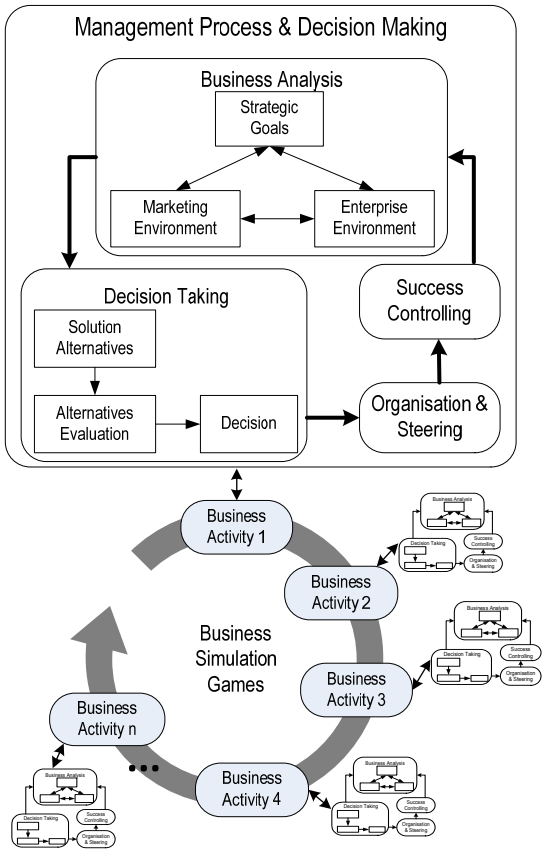


Figure 3: A conceptual framework of using DSS on top of business simulation games.

We will select business simulation games which support the representative business processes, evaluate a software platform for DSS provided SSBI functionalities and then attach with the business simulation games, lastly, create criteria for measuring the learning outcomes.

The framework will be used for experiments to prove that whether it is able to support learning and teaching BI, provide a better understanding of using SSBI tools and applications and SSBI technology will help end users to make better (valuable/actionable) decisions.

Furthermore, this framework will integrate an instrument for students and teachers to measure the learning outcomes based on the concept of "learning analytics" (Siemens, Gasevic, Haythornthwaite, Dawson, Shum, Ferguson, Duval, Verbert and Baker, 2011).

8 CONCLUSIONS

For many years, DSS have been used to improve the quality of managerial decisions. DSS applications have changed over the last decades, moving from Enterprise Reporting System to Management Information System and nowadays to Business Intelligence Solutions. The issue of teaching and learning DSS is still a big challenge in the academic world, since the DSS- or BI-related subjects are still difficult, complex and challenging. Moreover, the demand for well-educated students in the field of DSS is still growing.

We are working on a framework using business simulation games to overcome the restrictions and limitations of the existing DSS teaching solutions. We will embed a SSBI solution into business simulation games in order to learn and teach DSS/BI in a modern, integrated and fun-to-use environment.

We also believe that the integration of SSBI into business simulation games will increase the learning outcomes. We will provide a platform to measure and manage students' learning in the field of DSS/BI. Our platform will also be used for experiments to measure learning behaviour, with a strong focus on the 21st century skills defined by the European Community (Redecker, Leis, Leendertse, Punie, Gijsbers, Kirschner, Stoyanov and Hoogveld, 2011).

REFERENCES

Arnott, D. (2008). Personal decision support. systems. In Burstein, F. & Holsapple, C. W. (Eds.), *Handbook on decision support systems 2: Variations,* 127- 150. Berlin-Heidelberg: Springer-Verlag.

Ben-Zvi, T. (2010). The efficacy of business simulation. games in creating decision support systems: an experimental investigation. *Decision Support Systems and Electronic Commerce,* 49(1), 61-69.

Connolly, T. and Stansfield, M. (2006). Using games-based eLearning technologies in overcoming. difficulties in teaching information systems. *Journal of Information Technology Education,* 5, 459-476.

DSS 2.0 Conference. (2014). *"DSS2.0 – supporting decision making with new technologies"*, 2-5 June 2014, Universite Pierre et Marie Curie, Paris, France. Retrieved from http://dss20conference.wordpress.com/

Eckerson, W. (2012). The secrets of Self-Service BI. *Blog: Wayne Eckerson – BeyeNETWORK,* June 1, 2012. Retrieved from http://www.b-eye-network.com/blogs/ eckerson/archives/2011/01/the_secrets_of.php.

Evelson, B. (2012). Top 10 BI predictions for 2013 and beyond. Retrieved from http://blogs.forrester.com/ print/boris_evelson/12-12-12-top_10_bi_ predictions_ for_2013_and_beyond.

Faria, A. J., Hutchison, D., Wellington, W. J. and Gold, S. (2009). Developments in business gaming: a review of the past 40 years. *Simulation & Gaming,* 40(4), August 2009, 464-487.

Gluchowski, P., Gabriel, R. and Dittmar, C. (2008). *Management support systeme und business intelligence: Computergestützte Informationssysteme für Fach- und Führungskräfte.* Springer-Verlag Berlin Heidelberg.

Hackathorn, R. (2003). Minimizing action distance. *The Data Administration Newsletter*, July 1, 2003. Retrieved from http://www.tdan.com/print/5132.

Howson, C. (2013). 7 top business intelligence trends for 2013. Retrieved from http://www.information week.com/ software/information-management/7-top-business-intelligence-trends-for-2013/d/d-id/1108351?

Imhoff, C. and White, C. (2011). Self-Service Business Intelligence: empowering users to generate insights. *TDWI Best Practices Report.* Third Quarter 2011.

Inmon, W. H., Imhoff, C. and Sousa, R. (2001). Corporate information factory. 2nd edition. The United States of America: John Wiley & Sons, Inc.

Kulkarni, N. (2012). Embrace the future of BI: Self Service. *Information Management,* July 17, 2012. Retrieved from http://www.information-management. com/ newsletters/self-service-business-intelligence-bi-tdwi-kulkarni-10022855-1.html.

Lin, Y. L. and Tu, Y. Z. (2012). The values of college students in business simulation game: a means-end chain approach. *Computer & Education,* 58, 116-117.

Mundy, J. (2013). Design tip #153 three critical components for successful Self-Service BI. *Kimball Group,* March 4, 2013. Retrieved from http:// www.kimballgroup.com/2013/03/04/design-tip-153-three-critical-components-for-successful-self-service-bi/

Power, D. J. (2007). A brief history of decision support.

systems. *DSSResources.COM*. version 4.0, March 10, 2007. Retrieved from http://www.groupdecisionroom. nl/artikelen/decision-support-system.pdf.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S.B., Ferguson, R., Duval, E., Verbert, K. and Baker, R.S.J.d. (2011). Open Learning Analytics: an integrtated & modularized platform. *Society for Learning Analytics Research (SOLAR)*, July 28, 2011. Retreived from http://solaresearch.org/OpenLearning Analytics.pdf.

Shim, J. P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R. and Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems,* 33, 111-126.

Wawer, M., Miloz, M., Muryjas, P. and Rzemieniak, M. (2013). Business simulation games in forming of students' enterpreneurship. *International Journal of Economics and Management Sciences (IJEMS),* 3(1).

Williams, D. (2011). Impact of business simulation games in enterprise education. T*he 2010 University of Huddersfield Annual Learning and Teaching Conference.* University of Huddersfield, Huddersfield.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., Orvis, K. L. and Conkey, C. (2009). Relationships between attributes and learning outcomes: review and research proposals. simulation & gaming, 40(2), 217-266. doi:10.1177/1046878108321866.

Wixom, B. H. and Watson, H. J. (2010). The BI-based organization. *International Journal of Business Intelligence Research,* 1(1), 13-28.

Wixom, B. H., Ariyachandra, T. and Mooney, J. (2013). State of business intelligence in academia, *BI Congress 3,* 15-16 December 2012, Orlando, FL, USA.

Redecker, C., Leis, M., Leendertse, M., Punie, Y., Gijsbers, G., Kirschner, P., Stoyanov, S. and Hoogveld, B. (2011). The Future of Learning : Preparing for Change. *JRC Scientific and Technical Reports.*Retrieved from http://ftp.jrc.es/EURdoc/JRC 66836.pdf.

MOOC and Mechanized Grading

Christian Queinnec
Sorbonne Universités, UPMC Univ Paris 06, UMR 7606, LIP6, F-75005 Paris, France
christian.queinnec@lip6.fr

Keywords: MOOC, Mechanized Grading.

Abstract: As many others, we too are developping a Massive Online Open Course or MOOC. This MOOC will teach recursive programming to beginners and will heavily use an already existing infrastructure for mechanical grading (Queinnec, 2010). This position paper discusses how these two components are combined in order to increase students' involvement.

1 INTRODUCTION

Developping a MOOC is now a common activity in the Academia and so are we doing. This paper is a position paper that presents the main characteristics of our future MOOC: it makes an heavy use of an infrastructure to mechanically grade students' programs. How we intend to combine our MOOC with that infrastructure and how we want to create incentives for the students in order to increase their involvement is addressed in this position paper.

Programming exercises are proposed in between the videos of the course. These exercises ask for real programs from students. These programs are then mechanically graded and grading reports are sent back to the students. Reading grading reports allows students to evolve their programs but additional incentives may better help students. We propose to combine three means which should help students to progress by themselves.

- Students may program in pair: video-chatting around a common, shared program, that both students can edit, recreates conviviality. The underlying question is how to provide a good partner for a student who looks for one ?
- We also propose to students to peep slightly better programs from other students' and learn from them. The underlying question is how to select the most adequate programs to be peeped ?
- Finally, we ask student to recommend (or not) the peeped programs that were helpful, this ranking should help answering the previous item.

Section 2 presents the main lines of our MOOC and Section 3 presents the grading infrastructure while its new features appear in Section 4. Proposed incentives are described in Section 5 and Section 7 concludes this position paper.

2 PREPARATION OF A MOOC

We are currently developping a MOOC on recursive programming. The e-learning part is based on a course created in 2000 (Brygoo et al., 2002) and since then delivered every year at UPMC to hundreds of young scientific students as an introduction to Computer Science. The course material was, from 2000 to 2003, provided as a physical CDrom then, from 2004 to 2006, as a CDrom image (300MB) downloadable from UPMC web servers. An innovative characteristics of these CDroms was that they contained a programming environment (based on DrScheme (Felleisen et al., 1998)) with a local mechanized grader, see (Brygoo et al., 2002) for details. The students could then read the course documents, write programs, run them and be graded without requiring an Internet connection.

This new MOOC[1], adequately named "*Programmation récursive*", is an endeavour to extend this course to a broader French-speaking audience, to experiment with the aspects related to social networks and, finally, to collect and study students' answers to the proposed exercises in order to build appropriate error taxonomy and thus better future editions of the MOOC.

Of course, the context has dramatically evolved from 2000 to now. The mechanical grader of 2000

[1] http://programmation-recursive-1.appspot.com

which was grading Scheme programs (Scheme is the programming language used by that course) has now evolved into a multi-language grading infrastructure running in the cloud (Queinnec, 2010). The MOOC uses CourseBuilder (Google, 2013), is hosted on Google App Engine for elasticity, uses YouTube for video streaming and a Google group for forum.

The programming environment, named MrScheme is now provided as a Scheme interpreter, written in Javascript by Frédéric Peschanski and his colleagues (Peschanki, 2013). This programming environment runs locally in students' browsers. With help of MrScheme, students must test their programs (an habit we enforce) before requesting their program to be graded thus saving servers' computing power. We require students to write programs satisfying a given specification but also to write their own tests for their own programs. The machinery don't accept to grade programs that fail their own tests.

3 GRADING INFRASTRUCTURE

The grading infrastructure is named FW4EX and described in (Queinnec, 2010). This is a cloud-based infrastructure controlled by REST protocols. Teachers uploads exercises that are tar-gzipped sets of files containing scripts to grade students' programs. Students submit their programs and receive a grading report in return. The infrastructure also offers additional services such as the whole history of their submissions and their associated grading reports.

The infrastructure was carefully defined to scale up, see Figure 1. Students submit (1) to some acquisition servers which act as queueing servers that are regularly polled (2) by some marking drivers. As their name implies it, marking drivers grade student's submission and send (3) the resulting grading reports to long term storage servers (Amazon S3 for instance). Finally storage servers are polled (4) by the waiting students. Student's browsers choose the acquisition server which in return tells where to fetch the resulting future grading report.

Marking drivers are isolated, they cannot be queried. Moreover they run inner virtual machines to run the grading scripts. This allows to confine (potentially malicious) students' and teachers' programs in memory, time, cpu, access to Internet, etc.

Marking drivers also record (5), when connected, the details of the performed grading into a centralized database. Acquisition and storage servers don't perform heavy computations but there are more than one in order to offer some redundancy to ensure continuity of service. The number of marking drivers is handled elastically that is proportionally to the number of submissions to grade. To grade a submission more than once is not a problem: this is the price to pay to absorb grading peaks.

In 2011, we added a new service, see (Queinnec, 2011), that tries to rank students according to their skills. To submit a program is considered as a move in a game that this student plays against all other students who try the same exercise. If a student gets a higher mark in fewer attempts then the student wins over all other students who got a lower mark or a similar mark but in more attempts. Using a ranking algorithm inspired from Glicko (Glickman, 1995) or TrueSkill (Graepel et al., 2007), it is then possible to rank students on a scale bounded by two virtual students:

- the best student succeeds every exercise with the right answer at first attempt
- and the worst student fails every exercise with one more attempt than the worst real observed student.

This approach can only rank students having tried a sufficient number of exercises in order to appreciate their skill.

4 GRADING PROCESS

When a student submits a program, this program contains functions and their associated tests, let's call them f_s and t_s. For instance, the next snippet shows a student's submission for an exercise asking for the perimeter of a rectangle:

```
(define (perimeter height width)
  (* 2 (+ height width)) )

(check perimeter
   (perimeter 1 1) => 4
   (perimeter 1 3) => 8 )
```

This snippet contains the definition of the `perimeter` function followed by a `check` clause (an extension we made to the Scheme language) checking `perimeter` on two different inputs. The `check` clause implements unit testing. With other languages, Java for instance, we use the JUnit framework instead, (Beck and Gamma, 2012).

The author of the exercise (a teacher) has also written a similar program that is, a function f_t and some tests t_t. The grading infrastructure uses an instrumented Scheme interpreter and executes the following steps:

1. $t_s(f_s)$ should be correct. Student's program that fail student's own tests are rejected. This ensures

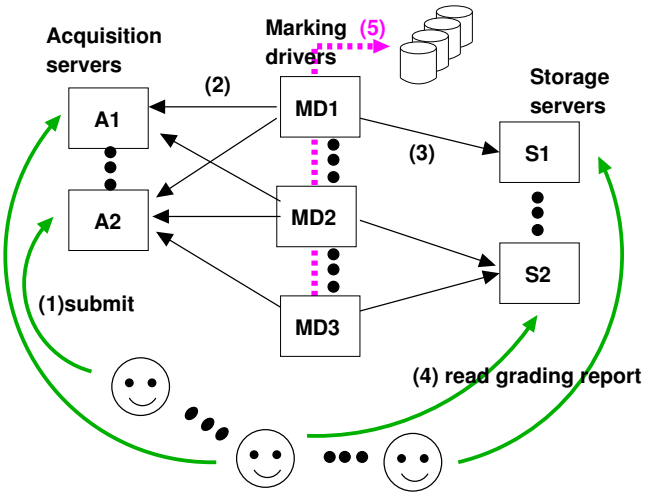


Figure 1: Architecture of FW4EX infrastructure.

that students don't forget to check their own code. We also check that the student does not obviously cheat i.e., t_s really calls function f_s.

2. $t_s(f_t)$ should be correct that is, student's tests should be related to the problem solved by f_t. The number of student's successful tests and the number of time the teacher's function has been called are used to provide a partial mark.
3. $t_t(f_s)$ should be correct that is, student's function should pass teacher's tests. The number of successfully passed tests also provides a partial mark.
4. The Scheme interpreter was instrumented in order to compute coverage profiles. Comparing the coverage of the student's tests with respect to teacher's test provides the last component of the final mark. Student's tests should at least execute all the code parts of their own code that are executed by teacher's tests. This again provides a partial mark.
5. All these partial marks are weighted and combined to form the final mark.

While steps 1, 2 and 3 were already present in the old CDroms, step 4 is new and measures the completeness of student's tests with respect to teacher's tests (which might be not perfect!).

Most exercises are graded in less than 15 seconds. The grading report returned to the student verbalizes what was submitted, which tests had been done and which kind of results were obtained with student's code compared to teacher's code. These reports are often lengthy but reading them carefully to understand the discrepancies develop students' debugging skill.

5 INCENTIVES

Equiped with such a grading machinery, we must offer incentives to the students so they may progress by themselves. Some incentives are currently under development and will be tested when the MOOC starts in 2014. The rest of the Section describes these incentives and the scientific challenges behind them which are not completely solved today.

5.1 Pair Programming

The first incentive is to provide an infrastructure for pair of students to work conjunctly on an exercise. This is pair programming as advocated by eXtreme Programming (Beck, 2000). A pairing server will be provided from which students will get a peer, the server will then provide a shared (Google) doc or equivalent and will let the two students work together and submit together. The server will choose a peer with roughly the same skill as determined by the ranking algorithm explained in Section 3. Of course, this might only work if a sufficient number of students in need of a peer are simultaneously present therefore, for every week of the MOOC, we intend to define peering periods. This feature will also require a widget in the shared doc to submit the joint work and see the resulting grading report. Accompanying this shared document with a Google hangout allowing to share voice and/or video will probably be attractive.

The quality of the peering process depends on the accuracy of the skill ranking. Conversely, to be able to appreciate students' skills allow the pairing server to also provide opponents with roughly the same skill:

the goal is then to propose, at the same time, the same exercise to the opponents and to compare their results in mark and time to reach that mark.

5.2 Epsilon-better Peeping

The second incentive is to propose to students having obtained a mark m and eager to progress, to peep two other students' submissions with slightly better marks i.e., $m+\varepsilon$. This is what we called "Epsilon-better peeping". This will favour reading other's code, another important skill worth stressing since beginners often think that they code for computers and not for humans! The slightly better mark may have been obtained by a better definition of the function or by better tests. Both allow to improve students' work. Reading these others' submission carefully and determine why they are better may be eye-opening.

To prevent students to just copy-paste better solutions, we will limit the number of times students may peep at other's submissions.

For students who stick to very low marks, we will probably have to set ε to a bigger value. If a huge number of students attend this MOOC, we may try various settings for this parameter to help students climb the first step.

5.3 Recommendation

After being served other's submissions, students will have to tell whether one of these other's submission was useful or not. This is a kind of recommendation system (or crowd ranking) from which the best helping submissions should emerge. However, differently from recommandation systems where a huge number of persons recommand a few items (movies for instance) here, we have a few students producing a huge number of submissions. Therefore to select the most appropriate submissions is a real challenge.

What we envision is to ask the teaching assistants to write a set of programs with increasing marks and, for the first edition of the MOOC, to favour these programs. This also solves the bootstrap problem since there must be other's submissions in order to implement this incentive.

Accumulating students' submissions should allow to elaborate a taxonomy of programs and errors. This taxonomy will help improving grading reports. Reports may include hints triggered by the kind of recognized error. The recommandation system that selects the best helping submissions, may also use that taxonomy. But this taxonomy will only be taken into account for the next edition of the MOOC.

6 RELATED WORK

Mechanical grading has been used for many years in many different contexts and programming languages. However generic architectures such as (Striewe et al., 2009), that support multiple languages, that are scalable and robust are not so common. Our infrastructure is one of them.

The estimation of students' skill is also a well studied domain (Heiner et al., 2004). Many works exist that try to characterize the student's model that is, its shape and its parameters (Jonsson et al., 2005) (Cen et al., 2006). They often start from an analysis relating exercises and the involved primitive skills then, they observe students' progress (mining the logs) in order to determine the parameters that best fit the model mainly with the "expectation maximization" technique (Ferguson, 2005).

The previous studies use far more information than us since they mine the logs of an intelligent tutor system where are recorded which exercise is delivered, how long the student read the stem, what help he requires, etc. By contrast, our grading infrastructure only gives us access to marks. Our set of proposed exercises is not (yet) related to the involved skills nor the set of skills is clearly stated. Therefore we are currently more interested to provide incentives to work in pairs with an attractive but rigorous feedback.

While recommendation systems are legion, to recommend the slightly better programs that helped to progress may be an interesting idea. We will see if our MOOC stands up to its promises.

7 FINAL REMARKS

In this paper, we present some ideas that are currently under development for a MOOC teaching recursive programming for beginners. This MOOC will start in March 2014 hence results are not yet known.

However and as far as we know, the conjunction of a grading machinery, a skill ranking algorithm and a recommendation system for help seems to be innovative and worth studying.

REFERENCES

Beck, K. (2000). eXtreme Programming. http://en.wikipedia.org/wiki/Extreme_programming.

Beck, K. and Gamma, E. (2012). The JUnit framework, v4.11. http://junit.org/.

Brygoo, A., Durand, T., Manoury, P., Queinnec, C., and Soria, M. (2002). Experiment around a training en-

gine. In *IFIP WCC 2002 – World Computer Congress*, Montréal (Canada). IFIP.

Cen, H., Koedinger, K., and Junker, B. (2006). Learning factors analysis - a general method for cognitive model evaluation and improvement. In *Paper presented at the 8th International Conference on Intelligent Tutoring Systems*, pages 164–175.

Felleisen, M., Findler, R., Flatt, M., and Krishnamurthi, S. (1998). The DrScheme Project: An Overview. *SIGPLAN Notices*, 33(6):17–23.

Ferguson, K. (2005). Improving intelligent tutoring systems: Using expectation maximization to learn student skill levels.

Glickman, M. (1995). The Glicko system. Technical report, Boston University. http://glicko.net/glicko.doc/glicko.html.

Google (2013). CourseBuilder. https://code.google.com/p/course-builder/.

Graepel, T., Herbrich, R., and Minka, T. (2007). TrueSkill^{TM}: A bayesian skill rating system. Technical report, Microsoft. http://research microsoft.com/en-us/projects/trueskill/.

Heiner, C., Beck, J., and Mostow, J. (2004). Lessons on using its data to answer educational research questions. In *Workshop Proceedings of ITS-2004*, pages 1–9.

Jonsson, A., Johns, J., Mehranian, H., Arroyo, I., Woolf, B., Barto, A., Fisher, D., and Mahadevan, S. (2005). Evaluating the feasibility of learning student models from data. In *Proceedings of the Workshop on Educational Data Mining at AAAI-2005*, pages 1–6. MIT/AAAI Press.

Peschanki, F. (2013). MrScheme. http://github.com/fredokun/mrscheme.

Queinnec, C. (2010). An infrastructure for mechanised grading. In *CSEDU 2010 – Proceedings of the second International Conference on Computer Supported Education*, volume 2, pages 37–45, Valencia, Spain.

Queinnec, C. (2011). Ranking students with help of mechanized grading. See http://hal.archives-ouvertes.fr/hal-00671884/.

Striewe, M., Balz, M., and Goedicke, M. (2009). In Cordeiro, J. A. M., Shishkov, B., Verbraeck, A., and Helfert, M., editors, *CSEDU (2)*, pages 54–61. INSTICC Press.

Integrating Computers, Science, and Mathematics
A Course for Future Mathematics Teachers

Alfinio Flores
Department of Mathematical Sciences, University of Delaware, Newark, DE, U.S.A.
alfinio@udel.edu

Keywords: Mathematics Teaching, Integration of Technology, Teacher Preparation, Geogebra, Python, Modeling, Robotics.

Abstract: A course for prospective secondary mathematics teachers was developed at the University of Delaware, based on professional recommendations to integrate science, technology, engineering, and mathematics in the preparation of teachers of mathematics. Students used GeoGebra, Cabri3D, and Mathematica to model phenomena in the physical, natural and social sciences. They used motion sensors and graphing calculators to study motion. They wrote Python programs to simulate random phenomena. They built a robot and controlled it with a computer program, and made explicit the mathematical and scientific concepts involved in the functioning of the robot. Several forms of formative and summative assessment were conducted during the course. Teachers learned alternative ways of looking at mathematical concepts, and established connections in mathematics and with other areas.

1 INTRODUCTION

Technological environments offer opportunities to students of different skills and levels of understanding to engage with mathematical tasks and activities (Hollebrands, 2007). However, the integration of technology in mathematics teaching has been slower than anticipated due to multiple difficulties teachers face (Hohenwarter, Hohenwarter, & Lavicza, 2008). Having the technology available in schools is not enough to guarantee that teachers will know how to use it to teach mathematics. Future and in-service teachers need to have opportunities to develop the expertise and know-how to be able to incorporate the use of technology in their own teaching (Lawless & Pellegrino, 2007).

The new course *Learning mathematics with technology* offered at the University of Delaware is geared for first-year prospective secondary mathematics teachers. The purpose of the course is to provide future mathematics teachers, early in their teacher preparation program, with the knowledge and experience of technology-based activities that foster mathematical communication, connections, reasoning, and reflection that help students develop better understanding in mathematics. The course integrates modern interactive technologies to emphasize the learning of concepts, problem solving, exploration in mathematics, mathematical modeling, and connections to physics and other sciences. The course is based on current research and theories about how students best learn mathematics (Boaler, 2008), incorporates professional recommendations on the preparation of teachers of mathematics and the type of mathematics they need to experience and learn themselves (Conference Board of the Mathematical Sciences, 2012; National Council of Teachers of Mathematics, 2000; Common Core State Standards, 2010, National Research Council, 2002, 2011) and incorporates best teaching practices. The use of writing to learn mathematics is an important component of the course.

2 CONTENT KNOWLEDGE FOR MATHEMATICS TEACHERS

This course emphasizes pedagogical content knowledge (Shulman, 1986, 1987) for mathematics teachers. The students learn new technology, connections of mathematics with science, and a richer set of connections among high school mathematical concepts. The course implements

strategies and methodologies to learn and teach mathematics with technology that could be used by future teachers in their own classrooms. Students actively participate in mathematical modeling of physical, biological, and social phenomena (Gordon & Gordon, 2010) using a variety of technology tools, such as TI CBR 2 motion sensors and GeoGebra (International GeoGebra Institute, 2013). They also use hands-on materials and the computer to further develop their own understanding of mathematics concepts, such as conics. Throughout the course, students use Python (Enthought, 2013) and other programming platforms to write their own short computer programs that involve the use of loops (FOR or WHILE) and the use of logical structures such as IF…ELSE.

How to work in cooperative groups was explicitly discussed several times during the semester. One students wrote in his end of class reflection: "At the beginning of class, when we talked about the concept of working together in a group, it helped me better understand how to truly work together and feed off of each other's ideas. Communication as they said, is the most important thing about it and that helped me truly understand the concept."

2.1 Motion Sensors, Velocity, and Acceleration

Students conducted several experiments with moving objects. They captured data for position vs. time with motion detectors and analyzed the corresponding graphs. The very first day they used constant-velocity vehicles (Printz, 2006). They followed with the 7 m jump of a small stuffed toy donkey hanging from a parachute. Students collected distance vs. time data and described how the different phases of the fall were reflected on the graph (before the parachute was completely deployed and after), and how the corresponding parts of the graph represented constant positive acceleration or constant velocity (initial accelerated motion followed by falling at the terminal velocity). One student wrote about the parachute activity on her end of class reflection: "The effect air resistance has on acceleration and velocity of a falling object was a concept today's lesson clarified for me."

Students also captured distance vs. time data for a bouncing basketball and analyzed the corresponding graph (Cory, 2010), to determine when the velocity was positive or negative, when was the velocity increasing or decreasing, when was it zero, etc. Especially challenging for them was to interpret what was happening to the velocity when the ball hit the ground.

2.2 Vanishing Points in a Rowing Competition

Students watched a video of a rowing competition (Allain, 2013) and then used GeoGebra and their knowledge of perspective and vanishing points (Figure 1 and Figure 2) to devise a way to determine who was winning the competition. Then they wrote a letter to a fictional television producer explaining the advantages of their method to make it easier for television viewers to see wheter a scull was moving faster than another.



Figure 1: Vanishing point at $t = 34$.



Figure 2: Vanishing point at $t = 38$.

2.3 Fitting Curves to Sets of Points

Students learned ways in which GeoGebra provides an alternative point of entry to the topic of fitting linear and quadratic curves to two or three points. First they played a game in pairs. One player would position two or three points on the screen, and the other player would use sliders corresponding to the coefficients of linear or quadratic functions and adjust the value of the slider to hit as many points as possible. Students expressed that using sliders allowed them to understand better the role of each of the coefficients, and the advantages and disadvantages of using different representations for quadratic functions, such as using the coordinates (h, k) of the vertex in the equation $y = a(x - h)^2 + k$ or using the general equation $y = ax^2 + bx + c$. In another activity students used sliders to

experimentally fit a parabola to a set of data representing annual growth of redwood trees vs. rainfall.

Later, students used sliders also to fit other types of curves such as exponential functions. Sliders were used in addition to analytical ways to determine the parameters of the exponential function. One activity was to fit an exponential model to the growth of the population in the United States during its first century.

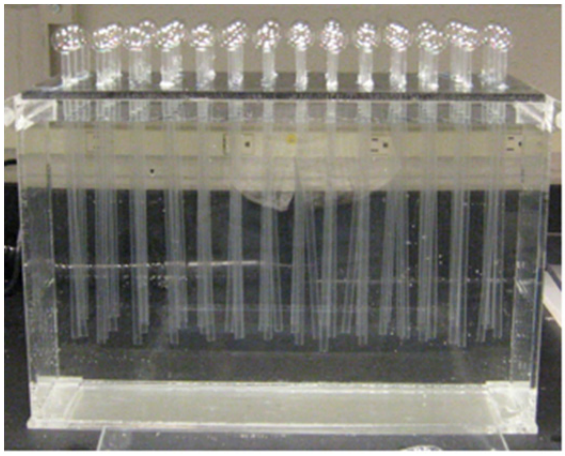


Figure 3: The Bubble Board.

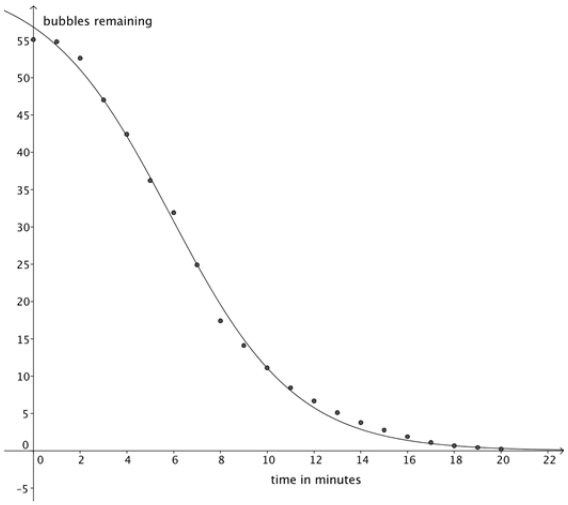


Figure 4: Fitting a logistic curve to the bubble data.

In another activity students generated 56 soap bubbles simultaneously with a bubble board (see Figure 3) (Hammons, Flores, Pelesko, Biehl, 2012) and recorded how many bubbles remained after intervals of one minute. Then they were given the data of the averages of nine such experiments and were asked to use GeoGebra to fit different mathematical models to the data (linear, quadratic, polynomial, exponential, and logistic) and discuss the advantages and disadvantages of each of the models. Students found that a logistic model

$$y = \frac{63.1}{1 + 0.11e^{0.37x}} \quad (1)$$

fit the data pretty well in the case of the bubbles (Figure 4).

2.4 Conics

In the United States often students' first encounter parabolas and other conics through equations such as $y = x^2$. Although textbooks mention that such curves can be obtained by slicing a cone, students usually do not have the opportunity to connect the different characterizations of conics. For example, for the ellipse, the curve can be described by the equation; using the locus definition (the sum of the distances to two fixed points is constant); by slicing a cone with a plane such that the angle of cut is bigger than the angle at the vertex of the cone; or in terms of the ratio of the distance from a point to a line and to a fixed point (focus-directrix). In this course students had the opportunity to establish connections among the different ways of characterizing conics using hands-on models, GeoGebra, and Cabri3D. One student expressed in her exit ticket: "Cabri3D helped me understand conic sections better. For the parabola, it helped me see that the distance to the directrix and focus is equal."

2.5 Python Programs to Simulate Random Phenomena

The instructor conducted several short sessions where a situation was presented, and then the class as a whole, using plain English, described the steps that needed to be included in a computer program to simulate the situation. For example, one die is tossed 1000 times. The program should keep track of how many Heads and Tails have occurred. After each toss the program determines who is in the lead and keeps track how many times each player has been in the lead. At the end the total number of outcome for heads and tails are displayed, and also how many times each player was in the lead, and how many times they were tied. The instructor then shared a program that would simulate the tossing of the die and keeping track of the outcomes in the desired way.

Students also did some thought experiments about situations involving random events and then

compared their results with those generated by a computer. For instance, students were asked to conduct a thought experiment of tossing a fair coin 100 times and fill a 10 by 10 table with the outcomes. Students thought simulations resulted in very fragmented tables, where tails and heads alternate frequently. Students were then given tables generated with *Mathematica* (Figure 5) to compare the distributions of the number of squares in the largest "island of randomness" (Flores, 2006) in their thought experiments and those generated by the computer.

H	H	T	H	T	T	T	H	H	H
T	H	T	T	T	T	H	H	H	H
H	H	T	T	T	T	H	T	T	T
T	T	T	T	T	H	H	H	T	H
H	H	T	H	T	T	T	T	T	T
H	H	T	T	T	T	T	T	T	T
H	T	H	H	H	T	T	T	H	H
T	H	T	H	H	H	T	H	T	T
H	T	H	H	H	T	H	T	T	T
T	T	T	H	T	H	T	H	T	T

Figure 5: Islands of randomness.

As one of their projects, students were asked to work individually or in pairs to write their own program to simulate a situation involving randomness. Two students wrote a program to simulate random sampling with replacement of marbles of three different colors (Figure 6).

```
import random

def MARBLE():
    T = 0
    L1 = []
    while T < 10:
        T = T + 1
        M = random.randrange(1,13)
        if M > 0 and M < 5:
            print "WHITE"
        else:
            if M > 4 and M < 9:
                print "RED"
            else:
                if M > 8 and M < 13:
                    print "BLUE"
        L1.append(M)
    print sorted(L1)
```

Figure 6: Random sampling with replacement.

2.6 A Robot and a Feedback Loop

The final project for the course is building a robot controlled by a program that includes a feedback loop. The instructor demonstrated in class a bumper car. The bumper car moves forward until it bumps into a wall. It moves back for one second, and turns for 1.5 seconds, and then repeats the cycle. Thus, inside the loop we want the bumper to do either of two things. Change direction and rotate if it bumps against an object, else keep moving forward if not.

Programs in Lego Mindstorms are constructed by dragging icons, rather than typing code. For example, the switch icon will open the IF… THEN… ELSE part of the program. If the touch sensor is pressed, the motor will execute the two command on the upper part (move in reverse, and turn). Else, the motor will continue moving forward (Figure 7).

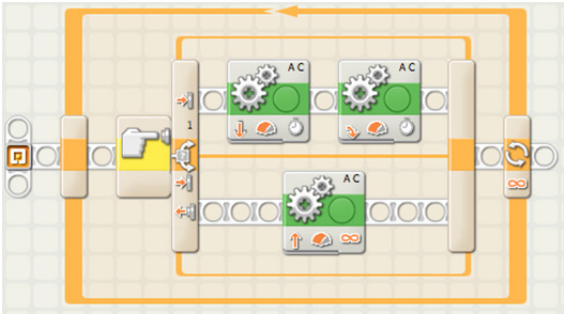


Figure 7: A program with a feedback loop.

One team of students chose to build a Robogator (Figure 8). Students described the structure of the program to control the Robogator in plain English:
IF an object is detected within 60 cm of the ultrasonic sensor
THEN the legs will move the robot forward infinitely
ELSE the robot does not move
IF the object is within 30 cm
THEN the jaws will open and close infinitely



Figure 8: Robogator.

Students also made explicit the scientific and mathematical concepts involved with the robot functioning: Velocity, speed, normal force, momentum, friction, centripetal acceleration, torque, feedback loops in nature depending on sensors. They

unpacked each one of these concepts further, for example,

"Torque: The Robogator is applying torque in order to move because its legs apply a force when rotating around an axis, which would be the leg joint to the motor. All of these legs are using the same amount of torque to move at a constant speed."

3 COURSE ASSESSMENT

Several forms of assessment were used throughout the course, both formative and summative, ranging from short answers in "exit tickets", informal verbal interviews, to a pre- and post-survey about conics, and a survey of the class as a whole. Students' learning was assessed through a midterm exam, written assignments, and their reports and in-class presentations of the final projects on Python programs and Lego robots.

For the most part, students found the use of technology for modeling exciting and illuminating. After the first session a student wrote on her reflection:

"It was surprising to begin experiments on the first day of class; it set a captivating tone for the rest of the semester. I was intrigued by the different approaches each pair of students took. I look forward to this class in order to improve my analysis skills. Hands-on activities are an extremely helpful tool in learning, and I am ready to have this class for my future teaching career."

At the end of the course, students were given a card where they wrote on one side what to keep in the course, and on the other what needed to be changed or deleted. Seven out the ten participants explicitly recommended to keep the use of GeoGebra.

Students participated actively in cooperative groups as they tackled complex tasks. Many had not worked in groups before in mathematics, and found the experience of learning from each other valuable. One student stated that the group work and collaboration with others "was good for learning how others think and helped me think about math."

Students found the course directly relevant for their future as secondary mathematics teachers. In addition, three peer tutors, more advanced mathematics education students, who participated in the course, also found this experience much more relevant than the required computer science course they took. Two of the peer tutors conducted an independent study related to the course. H. Kretz gave a survey about conics at the beginning of the course and then again after the topic had been covered. She found that in the beginning students described conics using only formulas and equations, with no mention of actually cutting a cone, or other geometric properties. For example, one of the students said "Each conic section is dependent on the equation." When asked how they would teach the topic of conics, most students described also approaches based on symbols, and only one student mentioned the use of computers as a tool for teaching conics. One student wrote "I would draw a conic on the white board and I would show a parabola on a calculator with $y = x^2$ and the textbook to try looking at different equations for conics". At the end of the course, their descriptions of the conics were more elaborated and they incorporated additional geometric elements and more connections. Likewise, at the end, the vast majority of the students mentioned the use of computers for teaching conics. A. Restrepo focused on students' attitudes and dispositions towards the group work, use of computers, and hands-on activities. After the midterm she found that 90% of the students thought that the hands-on activities helped them understand and recall better.

4 CONCLUSIONS

Our course was quite successful in introducing prospective teachers to alternative ways of looking at mathematical concepts, emphasize connections within mathematics, and with other areas. Van Voorst (1999) found in a course for in-service teachers that the use of interactive technology helped mathematics teachers see mathematics "less passively, as a set of procedures, and more actively as reasoning, exploring, solving problems, generating new information, and asking new questions" (p. 2). We found similar changes with the future teachers. Furthermore, our course illustrates a point emphasized in professional recommendations on the preparation of prospective teachers, that teachers should have the opportunity to experience content courses taught in the same way they are expected to teach in High School, courses that provide multiple points of entry to mathematical concepts through the use of computers, hands-on materials, and other tools. Students had also the opportunity to experience the advantages of working in cooperative groups and start developing their skills for making the group work more productive. The course will be taught again in the next academic cycle.

ACKNOWLEDGEMENTS

This course was developed as part of the grant *Integrated Science and Mathematics Education: A Model Course for Pre-Service Teachers* (S. Bernhardt, PI, A. Flores Co-PI, H. Shipman Co-PI) funded by the National Science Foundation. All opinions are of the author and do not necessarily represent the position of the funding agency.

REFERENCES

Allain, R. (2013). Physics of rowing. Retrieved December 28, 2013, from http://www.wired.com/wiredscience/2013/04/some-rowing-physics/

Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed ability approach. *British Educational Research Journal*, *34*(2), 167-194.

Common Core State Standards Initiative (2010). *Common Core State Standards for Mathematics*. Retrieved December 27, 2013, from http://www.corestandards.org/

Conference Board of the Mathematical Sciences (2012). *The mathematical education of teachers II*. Providence, RI: American Mathematical Society.

Cory, B. L. (2010). Bouncing balls and graphing derivatives. *Mathematics Teacher, 104*(3), 206-213.

Enthought (2013). *Canopy Python*. Retrieved December 28, 2013, from https://www.enthought.com/downloads/

Flores, A. (2006). Using graphing calculators to redress beliefs in the law of small numbers. In G. Burrill (Ed.), *Thinking and reasoning with data and chance* (pp. 291-304). Reston, VA: National Council of Teachers of Mathematics.

Gordon, S. P. and Gordon, F. S. (2010). *Functions, Data, and Models: An applied approach to college algebra*. Washington, DC: Mathematical Association of America.

Hammons, A. N., Flores, A., Pelesko, J. A., and Biehl, L. C. (2012). The "Bubble Board" and curve fitting. *Ohio Journal of School Mathematics*, No. 66, 9-16.

Hohenwarter, J., Hohenwarter, M., and Lavicza, Z. (2008). Introducing dynamic mathematics software to secondary school teachers: The case of GeoGebra. *Journal of Computers in Mathematics and Science Teaching*, *28*(2), 135-146.

Hollebrands, K. F. (2007). The role of a dynamic software program for geometry in the strategies high school mathematics students employ. *Journal for Research in Mathematics Education*, *38*(2), 164-192.

International GeoGebra Institute (2013). *GeoGebra*. Available from http://www.geogebra.org/cms/download.

Kretz, H. (2013). *Teaching mathematics with technology*. Unpublished independent study report. Newark, DE: University of Delaware.

Lawless, K. A. & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns and ways to pursue better questions and answers. *Review of Educational Research*, *77*(4), 575-614.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Research Council (2002). *Learning and understanding: Improving advanced study of mathematics and science in U.S. high schools*. Washington, DC: National Academy Press.

National Research Council (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: National Academies Press.

Printz, J. (2006). The buggy lab: Comparing displacement and time to derive constant velocity. *School Science and Mathematics*, *106*(5), 261-266.

Restrepo, A. (2013). *Research 2013*. Unpublished independent study report. Newark, DE: University of Delaware.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-22.

Van Voorst, C. (1999). Technology in mathematics teacher education. ICTE Educational Technology Resource Library. Retrieved February 4, 2014 from http://www.icte.org/t99_library/t99_54.pdf.

Learning Support System for Paleontological Environment Based on Body Experience and Sense of Immersion

Extinct Animals Move in Synchronization with Human Actions

Tomohiro Nakayama[1], Kaori Izuishi[2], Fusako Kusunoki[2], Ryuichi Yoshida[1], Takayuki Adachi[1], Takeki Ogitsu[1], Hiroshi Takemura[1], Hiroshi Mizoguchi[1] and Shigenori Inagaki[3]

[1]*Department of Mechanical Engineering, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba-ken, Japan*

[2]*Department of Computing, Tama Art University, Tokyo, Japan*

[3]*Graduate School of Human Development and Environment, Kobe University, Hyogo, Japan*

hmlab.tn12@gmail.com

Keywords: Kinect Sensor, Body Experience, FLASH Animation, Learning Support System.

Abstract: This paper proposes a simulation-based learning support system that animates extinct animals in synchronization with a human's location and actions. The system provides learners with a near-real body experience, and the sense that they have entered the paleontological environment. The aim of the proposed system is for this experience to improve the learning effect. Obviously, it is impossible for learners to experience the actual paleontological environment, and it is therefore difficult for them to understand the link between this environment and its actuality. This difficulty can be decreased by the proposed system, which provides learners with a real body experience and sense of immersion. At this point in time, the authors are implementing a system based on the proposed concept. As a preliminary evaluation, we conduct an experiment using the system, and interview the participants about the improvement in the learning effect. The experimental results confirm that most learners feel a tangible improvement in the learning effect due to the enjoyment of moving their body during the learning process.

1 INTRODUCTION

In recent years, environmental problems involving the destruction of nature have worsened. To improve our understanding of these problems, we must understand life in past geological periods and the biological environment of former generations. However, learners cannot watch or experience extinct animals and their environment. Thus, it is difficult to consider the learning target in relation to real things, which presents a difficulty in improving the learning effect.

Thus, to overcome this issue, we focus on the body experience and sense of immersion in learning. We are developing a more experiential simulation-based learning support system, in which the learner can enter the virtual environment of extinct animals. The system measures specific information, such as the human's location or actions. This allows the virtual environment and extinct animals to move across a large-scale screen. In this way, the system can compensate for near-real body experiences and immerse users in the environment. This experience and immersion lead to an improvement in the learning effect.

In this paper, as the first step toward realizing the above system, we describe a learning support system that moves in synchronization with the learner. After the learners have used the system, we interview them about the effectiveness of the proposed approach.

2 LEARNING THROUGH BODY EXPERIENCE

2.1 Introduction of Body Experience into Learning

Worsening environmental problems mean that education is becoming increasingly treasured as a way of relating the linkage between man and the environment. To develop an understanding of these problems, we must understand the wider transition of the biological environment from ancient to

modern times. However, although extinct animals and their environment actually existed, it is impossible for learners to watch and experience them. Additionally, it is impossible for learners to experience the wide-scale transition of the biological environment from such periods. There is a difficulty in learning about what one cannot actually experience, and so learners struggle to link their understanding of the animals and their environment to reality. Thus, learners may not realize that what they are being taught substantively existed, which prevents an improvement in the learning effect.

In this regard, we consider that the difficulty in immersing the user into the virtual environment is caused by a lack of body experience in the learning. The body experience of the learner is limited to watching or hearing the virtual environment from the outside, and so learners are not immersed in the environment(Akiko Deguchi, 2010; Takayuki Adachi 2012).

In recent years, it has become possible to replicate various environments with increasing degrees of reality using technologies such as computer graphics. Hence, the number of learning support systems that use these technologies in the form of an animation or game is increasing (El Rhalibi, 2011; Wernhuar Tarng, 2007; Michael Tscholl, 2013). These systems improve learners' motivation and the learning effect in a recognizable manner. However, learners can still only watch or learn from outside the virtual environment. Thus, they cannot obtain a satisfactory body experience and level of immersion, and it is difficult for them to consider what they learn from the virtual environment in relation to its substantive existence.

With the aim of improving the learning effect, we are developing an effective learning support system that permits a body experience (Figure 1). In this system, extinct animals and their environment are projected onto a large-scale display. Learners move the animals or environment, and obtain knowledge by moving their own body. To realize this, we use various sensors, and measure the learners' location, pose, and actions. Based on these measurements, the system can be controlled in a realistic manner. In this way, by changing the virtual environment according to body actions, learners feel more of a body experience and sense that they have entered into the environment. From this body experience, the system provides deeper immersion than watching illustrations or videos. This immersion can help to improve the learning effect.

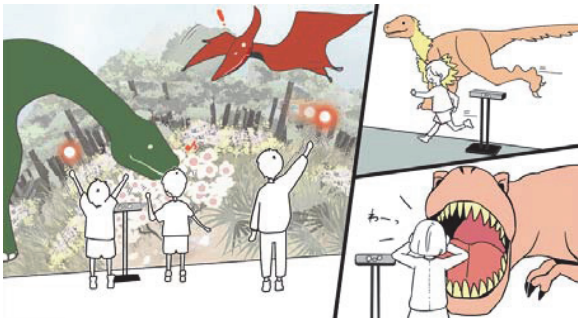


Figure 1: Proposed experiential learning support system.

Furthermore, moving the extinct animals in synchronization with the learner makes it possible to consider the extinct animal as being, in some sense, real, rather than imaginary. Thus, the learner can understand extinct animals more completely.

2.2 Current Implemented System

2.2.1 Overview

To implement the above learning support system, we are developing an animation system incorporating extinct animals. Figure 2 shows an overview of the system. This animation moves in synchronization with a human's location and movement. By playing with the content of this animation, the learner listens to information about the animal and their environment.

In this system, the extinct animal begins to walk when the human walks, and stops when the human stops. The background image scrolls according to the human's location. The learner sees the extinct animal following the pace of his/her walking. In this way, learners feel an affinity with the extinct animal, and obtain the sense that they are walking in the environment with the animal.

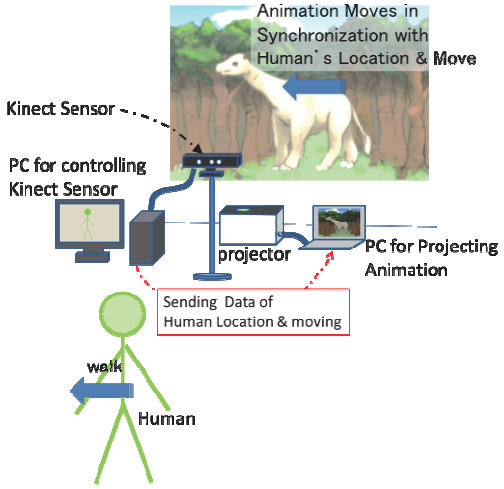


Figure 2: Overview of the -implemented system.

We expect that this experience and sense will lead to improvements in the learning effect.

2.2.2 Configuration of the System

The system must be able to control the animation in synchronization with the human's movement, which requires real-time knowledge of the human's location and movement status (walking or stopped). We utilize Microsoft's Kinect sensor for this purpose. Figure 3 shows the configuration of the system.

The Kinect sensor is a range image sensor, originally developed as a home video game device. Although cheap, the sensor can record advanced measurements of a user's location. Additionally, this sensor can recognize humans and the human skeleton using a library such as OpenNI. It can measure the location of human body parts, such as hands and legs, and we can detect the human's pose or status (e.g., walking, standing, sitting) from this function and location information(Jamie Shotton, 2011).

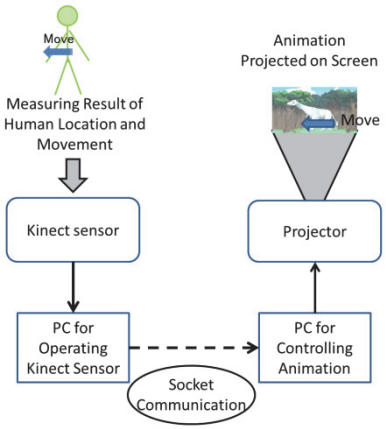


Figure 3: System configuration.



Figure 4: Measuring of human torso location.

We use these functions of the Kinect sensor and OpenNI library to recognize humans and detect the human skeleton. Then, we measure the location of the human torso (Figure 4). From the transition of pixel coordinates in the camera image of the human torso location, we can estimate the velocity and direction in which the human is walking.

The animation of extinct animals is created with ActionScript, and moves according to input numeric values. Information about the human's location, walking speed, or direction is sent to the PC that controls the animation. At this point, socket communication is used for the exchange of information. The PC controls the FLASH animation according to the information received, and the animation is projected onto the screen.

3 EXPERIMENT

3.1 Motion Experiment

First, we conducted a motion experiment to confirm that the animation system moves in synchronization with human walking.

As shown in Figure 5, we installed a Kinect sensor and a projector, and then asked a human to move across the front of the Kinect sensor once or twice, as shown in Figure 6. Thus, we can confirm that the following functions have been realized.

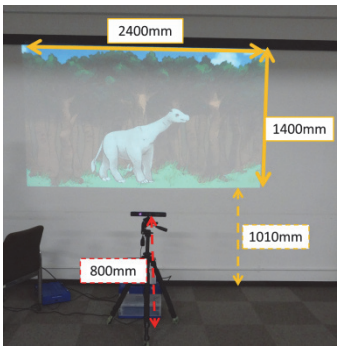


Figure 5: Installation status of sensor and projector.

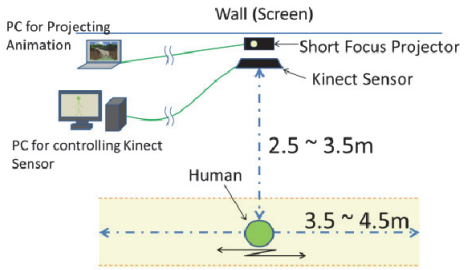


Figure 6: Movement of human in the experiment.

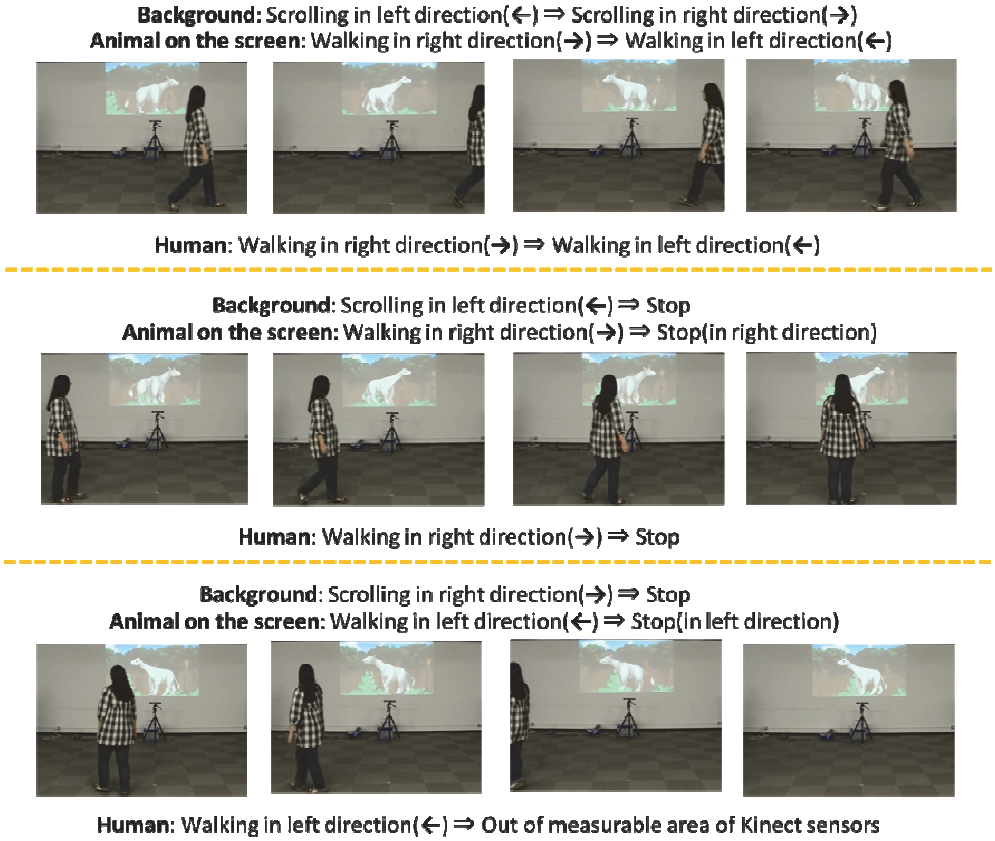


Figure 7: Experimental conditions.

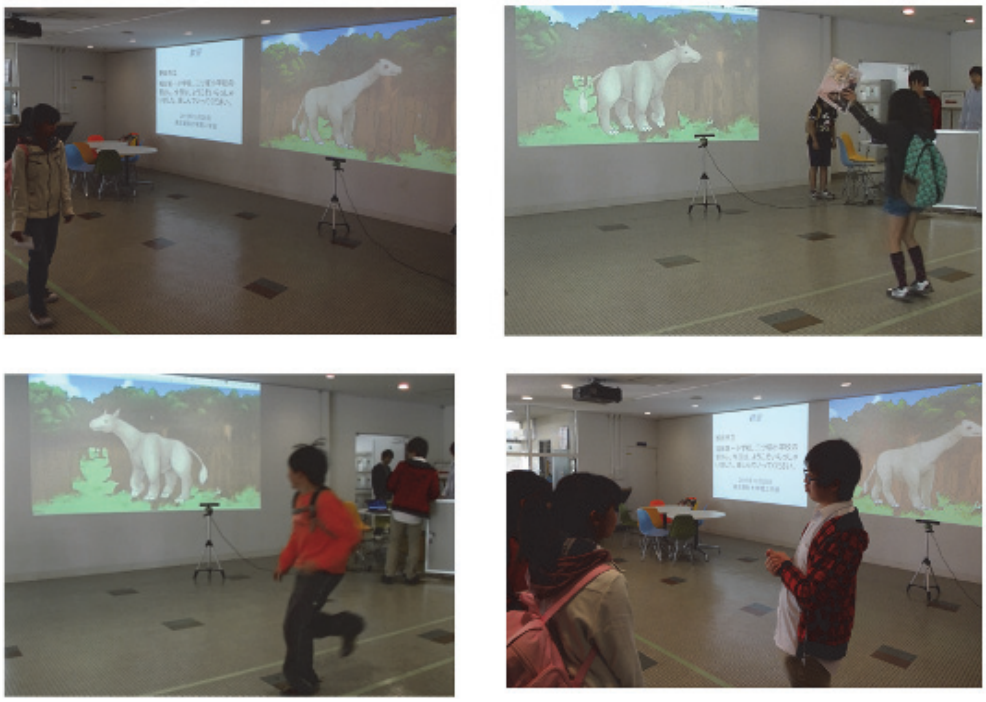


Figure 8: Evaluation experiment.

- Scrolling of background image linked with human location.
- Change in behavior of extinct animal on the screen according to human actions and direction of movement.
- Animation stops when the human moves beyond the measurable area of the Kinect sensor.

Figure 7 shows the experimental conditions. The animal on the screen moves in same direction as the human is walking, and stops when the human stops. When the human changes direction, the animal changes direction in synchronization with the human. Additionally, when the human moves out of the range of the Kinect sensor, the animal looks in the direction that the human walked away, as if watching after the human.

The background image scrolls in the opposite direction to that in which the human is walking, as if the human is moving forward in the environment. The background image scrolls according to the human's location, moving faster when the human walks faster.

From the above, we can confirm that we succeeded in realizing the proposed system in which animation contents move in accordance with a walking human.

3.2 Questionnaire Survey

Second, we conducted an interview to assess the improvement in the learning effect produced by the proposed system. We allowed children to play and learn with the system, and then conducted interviews with them.

The experiment employed 29 6th-grade children. As the children played with the proposed system, we explained about the extinct animal and environment on the screen (Figure 8). We then interviewed the children using the following questionnaire.

< Questionnaire>

Which learning contents help to progress your learning?

1. *This learning support system (proposed system)*
2. *Conventional learning method (e.g., textbook, classroom learning, model, movie, etc.)*
3. *Neither of the above*

Figure 9 shows the results of this questionnaire.

We then asked the children who chose "*1. Proposed system*" why they thought this was the case. Some of the responses are given below.

- "It is effective for enjoying learning to move my own body."
- "It is interesting that the extinct animal moves according to my actions."
- "We can remember the learning content, because it is fun."
- "It is very interesting that the animal follows that the animal follows my movements"

From these answers, we can confirm that the participants enjoyed learning by moving their bodies. By moving the animal and background according to their body actions, they felt a degree of enjoyment, and sensed that they were closer to the animal and its environment. In this way, most of the learners felt an improvement in the learning effect due to this sense of immersion.

Additionally, the following opinions were voiced:

- "It is interesting to learn the role and mechanism of the sensors."
- "I wonder how this system measures the human torso location."

Thus, the children took an interest in both the learning content and the system and sensors. From this, we believe the system led to an increased interest in intelligent machines.

One of the children who chose "*2. Conventional learning method*" stated, "In a museum, I feel the learning targets ancient things more deeply. So, in the case of studying old things such as extinct animals, I think it is better to study in a museum than with the proposed system." In a museum, many things are displayed that provide a sense of the ancient. In contrast, our system uses various modern devices, such as a PC and Kinect sensor. We think that seeing these devices may prevent users from 'feeling' the ancient period. Thus, seeing the system and its sensors may lead to an interest in intelligent machines, but these devices should be hidden so that learners cannot see them. Another opinion stated, "I have watched more large-scale exhibits. They are more dynamic and attractive than this system."

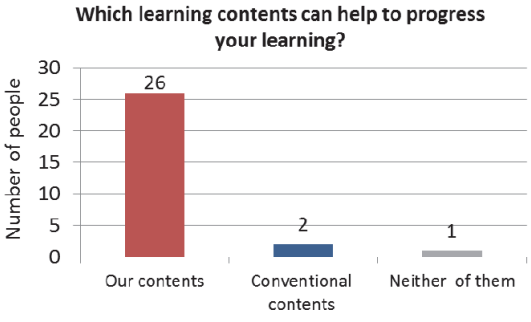


Figure 9: Questionnaire results.

At this point, we are using only one Kinect sensor and one projector. In future, we intend to increase the number of sensors and projectors, and

expand the scale of the system. Thus, we will make the proposed system more experiential than conventional learning methods or contents.

4 CONCLUSIONS

While learning about extinct animals and the environment in which they lived, learners cannot actually watch or experience the environment.

Hence, it is difficult for learners to understand the linkages between real things and attain a greater degree of learning. Thus, to compensate for this, we have focused on the body experience component of learning, and have proposed a learning support system in which users learn by moving their body.

In this system, the learning content is controlled according to the location, status, and actions of thelearner. By moving their own bodies, learners can obtain a near-real experience, as if they have entered the environment, and this improves their understanding of the learning target. In this paper, towards the realization of such a system, we developed a learning support system that moves in synchronization with the learner's location and actions (walking or stopping). After confirming that the system operated as intended, we conducted a questionnaire survey. From the results of the experiment, we have confirmed the potential for improvement in the learning effect due to the enjoyment of moving their bodies while learning. In the future, we intend to develop the system so that one can synchronize with not only walking also the other action. And then we intend to conduct evaluation test for more subjects.

ACKNOWLEDGEMENTS

I am particularly grateful for the illustration (Figure 1) provided by Ms Midori Aoki.

REFERENCES

El Rhalibi, A., Merabti, M., Ruwei Yun, Dan Liu, 2011, Game Based Learning Framework for Virtual 3D Dinosaurs Knowledge, *Developments in E-systems Engineering*, pages 419-424.

Wernhuar Tarng, Hsin-Hun Liou, 2007, The Development of a Virtual Dinosaur Museum, *Journal of Educational Technology Systems*, Volume 35, Number 4, pages 385-409.

Michael Tscholl, Robb Linfgren, Emily Johnson, 2013, Enacting Orbits: Refining the Design of a Full-Body Learning Simulation, *Interaction Design and Children 2013*, pages 451-454.

Akiko Deguchi, Shigenori Inagaki, Fusako Kusunoki, Etsuji Yamaguchi, Yoshiaki Takeda, Masanori Sugimoto, 2010, "Vegetation interaction game: Digital SUGOROKU of vegetation succession for children," *ICEC2010*, pages 493-495, 2010.

Takayuki Adachi, Masafumi Goseki, Keita Muratsu, Hiroshi Mizoguchi, Miki Namatame, Masanori Sugimoto, Fusako Kusunoki, Etsuji Yamaguchi, Shigenori Imagaki, Yoshiaki Takeda, 2013, Human SUGOROKU: Full-body Interaction System for Students to Learn Vegetation Succession, *Interaction Design and Children 2013*, pages 364-367.

Takayuki Adachi, Masafumi Goseki, Hiroshi Takemura, Hiroshi Mizoguchi, Fusako Kusunoki, Masanori Sugimoto, Etsuji Yamaguchi, Shigenori Inagaki, Yoshiaki Takeda, 2012,"Study on Identification of People and 3D Position Measurement with Ultrasonic Sensors and Kinect Sensors", *Proc. NOLTA2012*, pages 247-250.

Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake, 2011," Real-Time Human PoseRecognition in Parts from a Single Depth Image", *2011 IEEE International Conf. CVPR*, pages1297-1304.

Multiple Line System for Visualizing and Sharing Students' Commitment on the Tablet PC

Keita Muratsu[1], Hideo Funaoi[2], Etsuji Yamaguchi[1] and Shigenori Inagaki[1]
[1]*Graduate School of Human Development and Environment, Kobe University, 3-11 Tsurukabuto, Kobe, Japan*
[2]*Faculty of Education, Soka University, 1-236 Tangimachi Hachioji, Tokyo, Japan*
115d101d@stu.kobe-u.ac.jp, funaoi@umegumi.net,
etuji@opal.kobe-u.ac.jp, inagakis@kobe-u.ac.jp

Keywords: Visualizing Commitment, Sharing Commitment, Tablet PC.

Abstract: We enhanced and evaluated a system for visualizing and sharing the students' commitment to multiple ideas that runs on the iPad/iPad mini. This system has the following two functions: a) the function for the student to represent his/her commitment to multiple ideas, and b) the function to count the distribution of the students' commitment. We conducted a science lesson using this system targeting Japanese sixth grade elementary school students. Students were asked to input their commitment to the ideas each time they conducted an experiment. Subsequent to drawing the line the second time, the whole class discussed the commitment situations using the counting function. After the lesson, we investigated students' impression about the usability of the system and class discussion using the counting function. The results showed that this system was quite easy to use and effective in supporting learning that focuses on tracking the process of commitment change.

1 INTRODUCTION

It has been pointed out that it is important to visualize and share students' concept as an external representation in science education that support conceptual change (e.g., Duit and Treagust, 2012). This is because conceptual change is promoted when students share their own concepts and examine the differences among them.

Among the representation methods recommended by the theoretical study of conceptual change, a model that researchers and teachers consider promising is the conceptual ecology model by Posner et al. (1982). It regards conceptual change as a process of natural selection of multiple ideas (encompassing theory, knowledge, etc.) that exist within the intellectual environment. Each student is considered to have multiple ideas, rather than a single idea, ranging from the ones higher in the ecological hierarchy, which have successfully adapted to the intellectual environment, to those lower in the ecological hierarchy and facing extinction. A curriculum aiming to promote conceptual change in students by using the conceptual ecology model is required to visualize and share students' commitment to multiple ideas related to the lesson. In the past, a method that shows the commitment as a line by using pen and paper has been used (White and Gunstone, 1992).

In recent years, there have been attempts to display the commitment line on a system instead of using pen and paper. For example, Funaoi et al. (2006) and Nakashin et al. (2012) have developed systems to visualize and share the commitment to multiple ideas. These systems run on an iPod Touch. However, since the iPod Touch has an inherently small screen, its operation is not necessarily easy. In terms of visualization of commitment, difficulties have been observed in tracking the process of students' commitment change and its reason.

Therefore, in order to overcome the challenges described above, we developed a system in this study that visualizes and shares the commitment running on an iPad/iPad mini, which have become increasingly popular in recent years. The purpose of this study is to introduce this system to a science lesson at an elementary school in Japan, and conduct a preliminary assessment of its usability and effectiveness in supporting learning that focuses on tracking the process of commitment change.

2 OUTLINE OF THE SYSTEM

2.1 Development Environment

The development environment for the client side (iPad/iPad mini) is Xcode 5. The development environment for the server side is Windows 7 (Professional), Tomcat 6.0, and MySQL 5.5. The system can run on other operating systems if equivalent services are available.

2.2 Operation of the System

This System has the following two functions: a) the function for the student to represent his/her commitment to multiple ideas (hereinafter referred to as "the representing function"), and b) the function to count the distribution of the students' commitment (hereinafter referred to as "the counting function"). The representing function is performed on the iPad/iPad mini operated by each student. The counting function is performed on the PC operated by the teacher. The teacher's PC screen is displayed using a projector or large LCD TV in the classroom.

First, we will explain the representing function. Figure 1 shows the system configuration. Although it only shows one iPad/iPad mini, multiple devices can be connected to the server to share the inputted commitment. Once the System is launched and the school, grade, class, group, lesson unit, name of the student, and password are entered, the commitment input screen is displayed as shown in Figure 2. The vertical axis represents the level of commitment and the horizontal axis represents the progress of the lesson in time. First, tapping the button on the bottom left corner of the screen shown in Figure 2 will allow you to begin inputting your commitment. You can specify the level of commitment by dragging the square cursor up and down, and then confirm the level by tapping the cursor. Once confirmed, a line connecting the level of commitment inputted earlier to the level of commitment inputted most recently is displayed. When inputting commitments for different ideas,

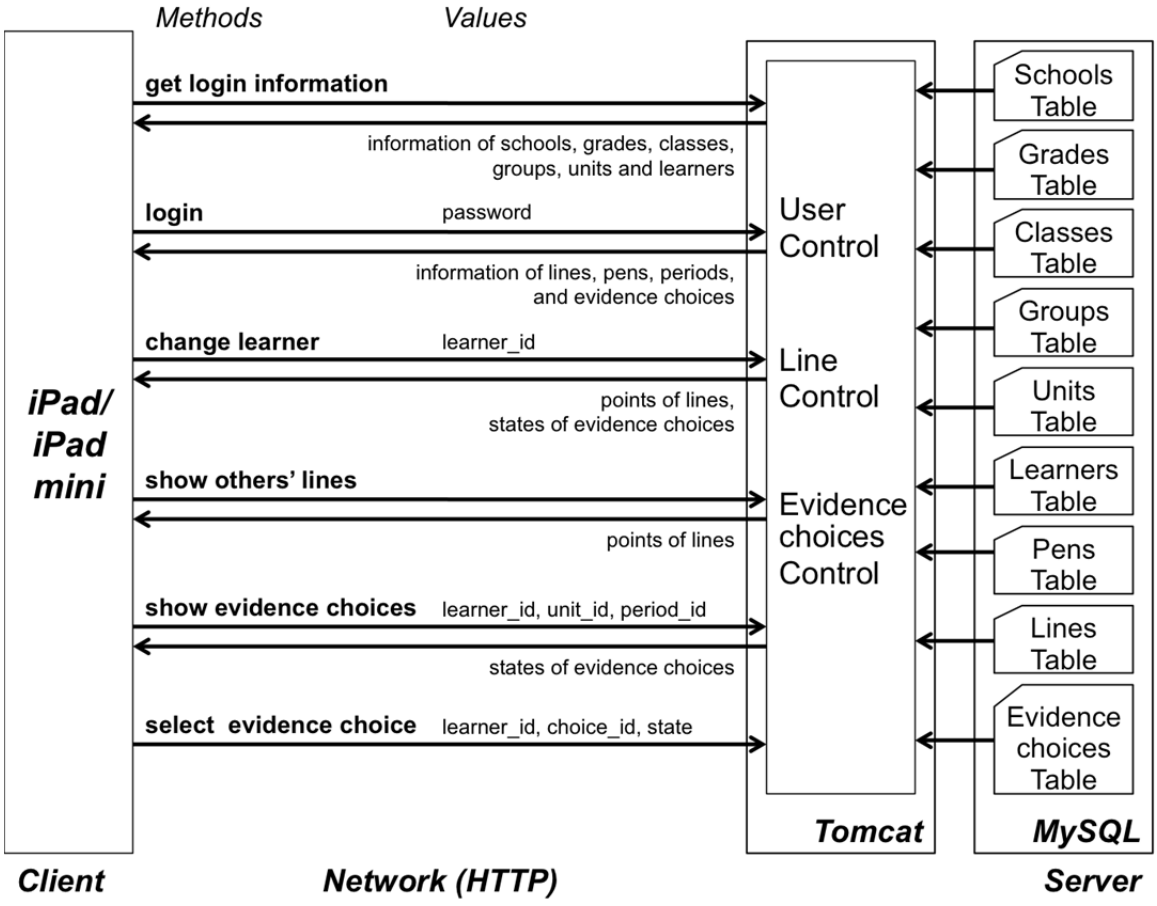


Figure 1: The system configuration.

tapping the button on the bottom right of the screen shown in Figure 2 will display the idea selection screen for inputting the commitment. The data on inputted commitment is sent to the server via the network and then stored in a MySQL Database through the control unit in Tomcat.

Next, we will explain the counting function. The counting function features the ability to list combinations of commitment changes to multiple ideas. When the request is sent from a teacher's PC, the commitment information inputted by students is read from the MySQL Database and then the combination of commitment changes to multiple ideas is displayed. This counting function focuses on the commitment inputted by students for the second time among all instances of input. It is possible to count the number of students who changed their commitment in the second round of input from "correct" to "incorrect", or vice versa and display a matrix of the commitment change combinations for two ideas. Figure 3 is the counting function screen. Here we can see that there were three students who changed their commitment to Idea A from "incorrect" to "correct", and their commitment to Idea B from "correct" to "incorrect". The use of this function allows us to track the process of the change in commitment to two ideas.

3 EVALUATION OF THE SYSTEM

3.1 Evaluation Method

3.1.1 Purpose

The purpose of the evaluation was to clarify the following two points: 1) whether the system is easy for the participants to use, and 2) whether class discussion using the counting function promotes the sharing of commitment among students.

3.1.2 Participants

The participants included 40 sixth-grade students (11-12 years old) at an elementary school in Japan. They had never used this System prior to this study. One iPad mini was assigned to every two students for inputting their commitment. The counting function was performed on a laptop operated by the teacher. The teacher, who was certified to teach

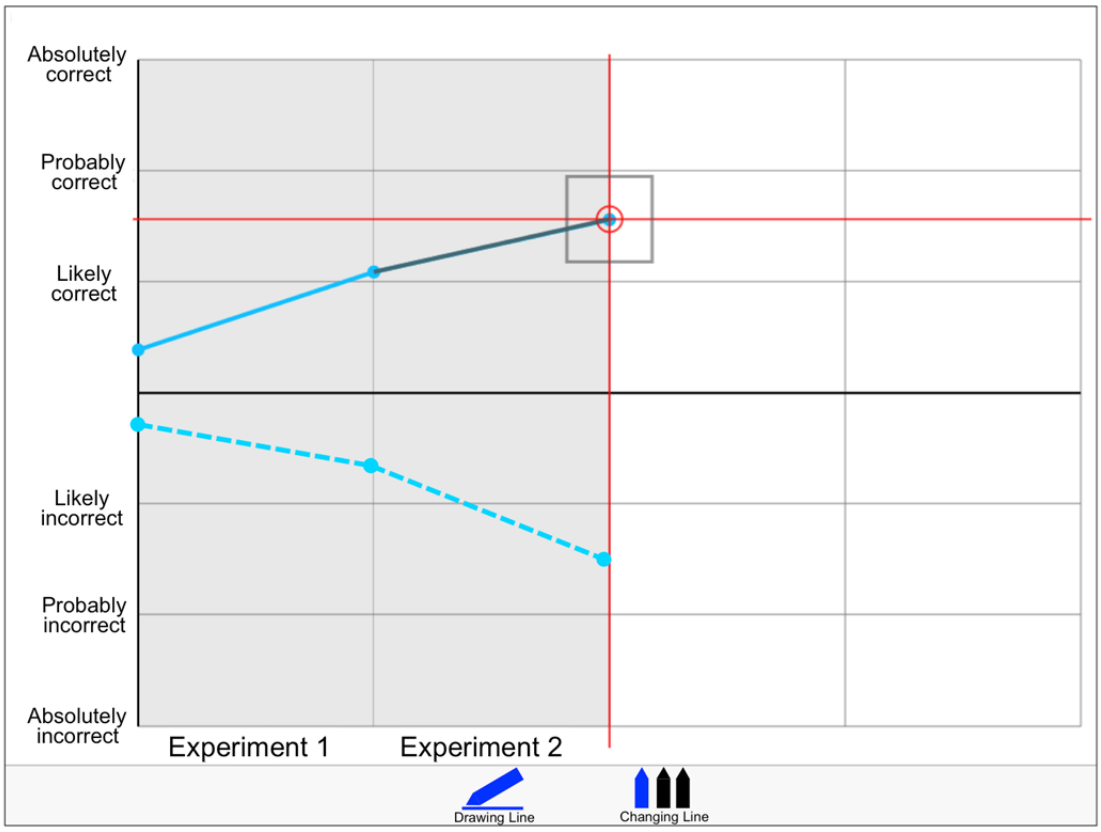


Figure 2: The commitment input screen.

science, had 15 years of teaching experience. The PC screen was mirrored on a LCD.

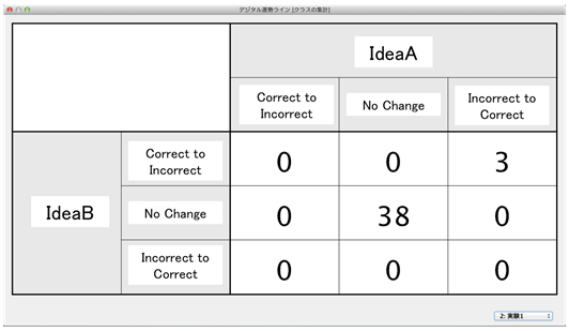


Figure 3: The counting function screen.

3.1.3 Tasks

The evaluation tasks were to investigate the students' impressions about the usability of the system and class discussion using the counting function.

3.1.4 Procedure

First, a total of nine science curriculum units (one unit is 45 minutes) were conducted. The lesson subject was static electricity. The aim of the lesson was to change a scientifically invalid idea ("There are two types of non-conductive objects: those that become electrically charged, and those that do not," referred to by the teacher and the students as the "Two Types Theory") to a scientifically valid one ("All non-conductive objects become electrically charged," which was referred to as the "All Theory") through hands-on experiments and discussions. Students were asked to input their commitment to these two ideas each time they conducted an experiment. In the first experiment, the students brought a charged drinking straw, charged paper bag, and charged plastic ruler into close proximity to a charged drinking straw. In the second experiment, they brought several charged objects into close proximity with a thick enamel wire. The charged objects were the following: a drinking straw, plastic ruler, candle, glass rod, and rubber hose. In the third experiment, they brought a charged drinking straw into close proximity with several non-charged objects: a piece of wire, plastic ruler, glass rod, spring onion, and pretzel snack. The line was drawn for a total of four times during the science curricular units. Subsequent to drawing the line the second time, the counting function screen was displayed on a large LCD every time the line was drawn, and the whole class discussed the commitment situations. After the lesson, an interview survey was conducted among four randomly selected students regarding the evaluation tasks. The tasks were performed using the individual interview method, which took about 10 minutes per person.

3.2 Results

First, we will describe the results of the interview survey on the usability of the system. Table 1 shows comments regarding the usability of the system extracted from the interview survey. S1 stated that it was possible to quickly draw a line when inputting a commitment because the drag and drop response time was good. This is a comment related to the response speed of the System. S2 commented on the screen size and the ease of the steps used to draw a line. S2 said it was easy to draw lines because there were no unnecessary steps before inputting the commitment. Regarding the iPad mini's screen size, S2 said it was appropriate and pointed out that it could become difficult to read if the screen size were smaller than that of iPad mini. Both S3 and S4 mentioned the fact that they were able to get used to inputting the commitment. They reported it was because they were able to immediately grasp the information necessary for operation.

Next, we will describe the interview results about the counting function. Table 2 shows comments regarding the counting function. First of all, S1 mentioned a case in which he asked why a few students changed the commitment that they supported. S1 said he understood from the counting function that there were three students who changed their commitment to All Theory from "incorrect" to "correct;" consequently, he became interested in those three students since only a few in the class made this commitment change, and felt like asking why they changed their commitment. Next, S2 stated that from the counting function he understood that the majority of students did not change their commitment to All Theory and Two Types Theory, which made him want to ask each student why their commitment didn't change. This is a comment on a case in which a student asked why the commitment supported by the majority didn't change. Furthermore, S3 mentioned a case in which he asked others who made a commitment change different from his to express their reason. S3 did not change his commitment because he thought All Theory was incorrect at the first opportunity he had to input his commitment as well the second. However, he said that it made him want to ask why there were others who changed their commitment to "correct" in the second round even though they had thought All

Theory was incorrect, as S3 did, when the first commitment was inputted. Finally, S4 said that he became eager to hear what others thought when he noticed there were others who changed the commitment as he did, because he thought there might be other reasons even though the commitment change was the same. This is a comment about a case in which a student asked why others made the same commitment change.

Table 1: Comments regarding the usability of the system, extracted from the interview survey.

S1:	It was nice because when I used the system to draw a line, it responded well to normal touch and I was able to draw the line in a minute or so.
S2:	It was easy to use because it is compact, there are no unnecessary steps to draw a line, and the procedure to draw a line was easy. The size of the screen was also just right. I think it would probably be difficult to read the screen if it were smaller than that.
S3:	Using this system was really easy, and I immediately found out what was written where. I think I also got used to writing on it right away.
S4:	I got used to the steps to draw a line within minutes because it was easy to draw the line. The screen was okay to read.

S1:Subject1 / S2:Subject2 / S3:Subject3 / S4:Subject4

Table 2: Comments regarding the counting function, extracted from the interview survey.

S1:	Even though there were only three who changed from “incorrect” to “correct”, I think it was rather useful because it got me interested in how people changed their thinking based on this experiment.
S2:	Because I immediately saw on the screen that there were many people who did not change, I thought it would be a good idea to ask each person what he/she thought.
S3:	I kept thinking that All theory was incorrect. So when I saw there were three people who changed from “incorrect” to “correct”, it made me wonder why they changed their mind, why they thought All theory was correct, and so on. It just got me thinking.
S4:	It made me want to ask people who changed their mind from All theory is “incorrect” to “correct” their opinions. I wanted to ask because I changed my mind the same way, you know, and I thought it might influence my thinking if I heard the reason why they thought that way.

S1:Subject1 / S2:Subject2 / S3:Subject3 / S4:Subject4

4 CONCLUSIONS AND FUTURE WORK

The purpose of this study was to introduce the system for visualizing and sharing the students’ commitment running on the iPad/iPad mini to a science lesson at an elementary school in Japan, and to conduct a preliminary assessment of its usability and effectiveness in supporting the learning.

First, we will consider the interview results regarding the usability. The results of the survey showed the operational effectiveness from the following four perspectives: 1) the responsiveness of the screen when inputting commitment is fast, 2) the steps to input commitment are easy, 3) the information necessary for inputting commitment is easy to understand, and 4) visually confirming information such as words on the screen is easy because the screen is large. Based on these results, we found that the System was easy for the participants to use.

Next, we will consider the interview results regarding the counting function. The four cases obtained in the interview survey can be roughly divided into two categories. The first are cases in which a student finds other individuals whom he wants to ask questions based on the population distribution trends displayed by the counting function, as in the cases of S1 and S2. There, the students were interested in the population distribution that stood out, such as commitment changes supported by a few students or by the majority. Therefore, the examination of the difference in ideas necessary for students’ conceptual change may be promoted. The second are cases in which, like S3 and S4, a student compares their own commitment change to other individuals’ commitment changes and finds whom to ask questions. There, students were interested in others who made the same commitment change or a different commitment change. This case may be more desirable because it may promote the examination of the differences in ideas, which is necessary for conceptual change among students, even more so than in the case of S1 and S2, by allowing students to compare commitment changes among others with their own commitment change for examination.

The future tasks include the following two. The first is to conduct a quantitative analysis of the usability of the system and the effectiveness of the counting function in supporting the learning. The second is to qualitatively analyze comments made by students and the teacher during the class discussion that used the counting function.

ACKNOWLEDGEMENTS

This research was supported by JSPS KAKENHI Grant Number 24300270.

REFERENCES

Duit, R., & Treagust, D. F., 2012. How can conceptual change contribute to theory and practice in science education? In Fraser, B. J., Tobin, K., & McRobbie, C. J. (Eds.), *Second International Handbook of Science Education,* 107-118. Dordrecht, The Netherlands: Springer.

Funaoi, H., Misawa, N., Deguchi, A., Inagaki, S., & Yamaguchi, E., 2006. Digital Fortune Line System: For externalizing and sharing learners' commitments to learning material. *Proceedings of E-Learn 2006*, 2823-2828. Honolulu, Hawaii.

Nakashin, S., Yamaguchi, E., Funaoi, H., Murakami, M., & Inagaki, S., 2012. Enhancement and evaluation of the counting function of the digital fortune line system toward supporting science learning. *Proceedings of E-Learn 2012*, 2044-2049. Montreal, Canada.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A., 1982. Accommodation of a scientific conception: Towards a theory of conceptual change. *Science Education*, 66 (2), 211-227.

White, R., & Gustone, R., 1992. *Probing Understanding*. London: The Falmer Press.

The Virtual Classroom
A Pilot Case in Inquiry Based Learning

Pavel Boytchev, Eliza Stefanova, Nikolina Nikolova and Krassen Stefanov
Faculty of Mathematics and Informatics, Sofia University, James Bourchier blvd, Sofia, Bulgaria
{boytchev, eliza, nnikolova, krassen}@fmi.uni-sofia.bg

Keywords: Virtual Classroom, Inquiry Based Learning, Inquiry based Science Education, WeSPOT, Educational Simulation.

Abstract: This paper discusses a pilot case held in First Private Mathematical School, Sofia, Bulgaria that explores the Inquiry Based Learning through a specially developed virtual classroom. The paper describes the motivation for using Inquiry Based Learning and how it will be gradually implemented through several pilot experiments. It reveals the design and the educational concepts embedded in the software tool used in the pilots: virtual classroom. The paper concludes with a discussion about the results of the pilot and plans for the next round of experiments.

1 INTRODUCTION

Modern conception of learning (Brown and Adler, 2008; Kolb, 1984) presents the acquisition of new knowledge and skills as a result of social interactions and a practical solution of problems and tasks. Students interact with objects of the reality, formulate their statements and assumptions, and seek to justify, prove, or refute them. Unfortunately, the practice in most educational institutions does not comply with the requirements of the theory. Students at secondary schools and universities are mostly in a passive role in the classroom, and teachers are often in the role of mentors.

One approach to solve the problem with the gap between theory and practice is the inquiry-based science education (IBSE) approach, in which students play the role of explorers and scientists as they try to address issues set by themselves, while finding answers to these questions is challenged by their own curiosity. This approach leverages a meaningful context for students to learn concepts by linking them with their personal experiences and insights. It leads to collecting structured knowledge in the given field of education and development of skills to carry out effective research.

The project weSPOT (2012) aims at supporting the implementation of this approach through the design, development and testing of appropriate software tools that will enable students to:

- Customize their environment for IBSE;
- Build, share and research either individually or in collaboration with their peers.

Thus, weSPOT aims to enable the connection of everyday life with training in subjects related to natural sciences in schools through the use of ICT (Mikroyannidis et al, 2012).

From the perspective of European teachers, weSPOT project will enable both teachers and students to apply an inquiry learning approach based on experiments carried out in a real school environment. Such experiments can be supported by computer simulations, 3D animations and video allowing students to understand better the subject of natural sciences.

This paper presents the creation of one such tool and its application in a weSPOT pilot.

2 weSPOT PROJECT

2.1 weSPOT IBL Model

A new Inquiry-Based Learning (IBL) model is developed in the weSPOT project framework. It follows the real reseach cycle with emphasis to develop students' inquiry metaskills (Mikroyannidis et al, 2013). It shares many of the phases described by Mulholland et al (2012), but it is more detailed in the description of things that teachers and students

should consider when doing inquiry.

The main phases of the weSPOT IBL model are:

- Question/hypothesis - students/learners decide on a topic or area of interest and try to formulate the questions or hypotheses that would like to pursue.
- Operationalisation - refers to the realisation of an idea with an aim to measure.
- Data collection - testing a hypothesis and seeing whether the real world behaves as predicted by the hypothesis.
- Data analysis - process of inspecting, cleaning, transforming and modelling data with the goal of highlighting useful information, suggesting conclusions and supporting decision making.
- Interpretation/discussion - describes the relevance of the results in relation to the question or hypothesis.
- Communication – reflection, sharing results.

2.2 Pilot Inquiry Scenario

Following the IBL model a pilot scenario "My classroom – energy efficient" was developed under the testbed "Energy effective buildings".

The scenario challenges the students with the question "How to make my classroom more energy efficient?"

During the Question/hypothesis phase the students should number factors influencing the energy consumption in the classroom (outside temperature, location of the room, heating, isolation, etc.). At the next stage they should decide what and how to measure and what additional data to collect to prove or reject their hypothesis. After data analysis and its interpretation the students should come to conclusions and present recommendations to the school principle. Their opinion should be scientifically proved by the elaborated data through public presentation in front of the school managers, parents, classmates and other stakeholders.

2.3 First Pilot Results

Three consecutive pilot runs of the scenario "My classroom – energy efficient" are planned. The first pilot during the spring of 2013 was used to validate the model. In this paper we describe the second pilot run in the autumn of 2013, in which the new software tool: Virtual Classroom simulation was used.

The idea to develop Virtual Classroom simulation was born during the first weSPOT pilot experiment (Stamenov and Dimitrova, 2013) in the First Private Mathematical School. The pilot was



Figure 1: A student measuring the outside temperature.

delivered in three classes in parallel with 6th graders in the frame of the subject Human and Nature. Its primary goal was to pilot the weSPOT IBL model.

The students' research task was to figure out which factors influence the consumption of electricity in the classroom and to propose measures to preserve it. As a part of the pilot, students had to collect data for outside and inside temperature of their classrooms. In order to do this they had measured several months, three times daily, these temperatures and wrote down them in order to use data later for their inquiry (Figure 1).

The observations and conclusions from the first pilot are very encouraging. They are used to give feedback to the weSPOT researchers and developers.

One of the conclusions gathered from the pilot was that for such young students (12-13 years old) several months inquiry is too long and exhausting, leading to loose of motivation. That is why it would be perfect if it is possible to speed up data collection. For the 2013/2014 academic year it was suggested to develop software simulation. The benefits of such computer program are:

- Time and duration of the pilot is independent of the particular season: it is possible the experiments



Figure 2: The initial scene of the virtual classroom.

to start and finish at any time of the year and, in the same time, the students can explore all seasons.
- The experiment is repeatable and the dataset is reusable: it is possible to set up the same situation with the same input data and to observe if the output is the same and if the hypothesis is proven again.

3 THE VIRTUAL CLASSROOM

The core technological component developed for the pilot case is a specially designed virtual classroom. The software presents a room with windows and a door. The room is in an environment with specific weather conditions. There are several devices like thermometers, a wall clock, and air conditioners. Users may navigate within the scene examining different part of it. Figure 2 shows the initial screen of the virtual classroom.

Figure 3 shows the general architecture of the virtual classroom. A set of external factors (sun, moon, clouds, rain and snow) affect the external air temperature. This temperature together with the internal factors (door, windows, air conditioners and students) and the elapsed time determines the temperature inside the room.

There are three virtually analogue devices (two thermometers and a clock) that are both input and output devices. They provide a way for the students to observe and measure temperatures and time, as well as they allow the student to setup manually the temperatures and the time.

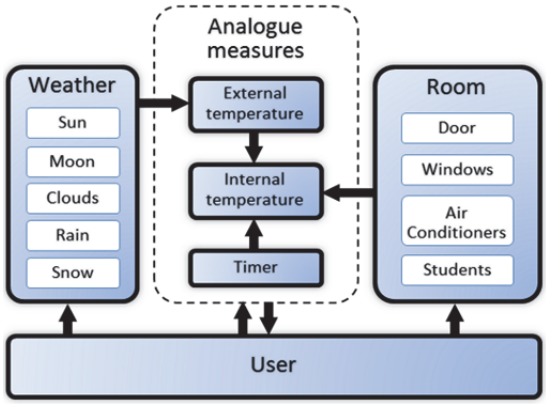


Figure 3: Simulation architecture.

The following subsections describe some basic concepts of the virtual classroom. Some of them refer to the functional design of the simulation (like weather condition, room configuration, thermal and temporal control), others refer to the educational design.

3.1 Functional Design

3.1.1 Weather

The virtual room model provides functionality to set up specific weather conditions, which are used to represent various inquiry scenarios.

There are two general modes, day and night, that are represented by a change of the light and the presence of an image of the Sun or the Moon. The weather conditions can be sequentially selected. The possible variations are to add clouds, rain or snow.

The modes together with the weather variations generate 10 combinations, which provide enough diversity of the external conditions. Each of these combinations has its own specific impact on the changes of the external environment. For example, if the outside temperature is 10°C and the users activate snowing, then the temperature will gradually drop down.

3.1.2 Room

The class room is the main object in the virtual room simulation. It can be modified by the user. There are four main elements that can be controlled – the door, the windows, the students and the air conditioners.

The door and the windows contribute to the overall openness of the room. The door can be either open or closed, while the windows have several levels of openness, randomized within preset ranges. Each window can be completely or partially opened (either vertically or horizontally). Figure 4 displays the appearance of completely opened windows.

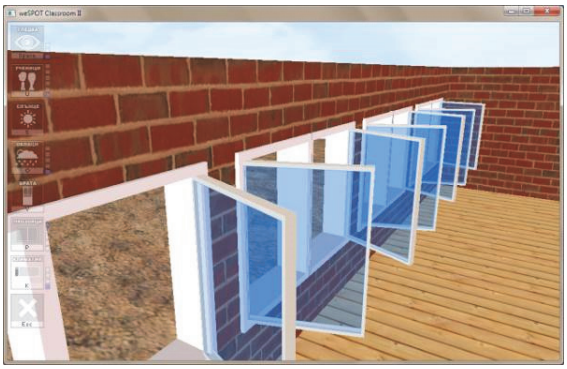


Figure 4: Fully opened windows.

The third controllable element in the classroom is the number of students. There are six predefined configurations ranging from an empty classroom (0 students) to an overcrowded one (40 students)

shown in Figure 5.

Finally, there are two air conditioners in the room. Initially they are turned off. Users can switch them on to cooling mode (18°C), room temperature mode (22°C) or heating mode (26°C).

3.1.3 Measures

The options for weather conditions and room configurations provide somewhat indirect control on the temperature inside the room.

The room, however, is equipped with two thermometers. One of the thermometers is attached to the outside of the room and measures the external temperature. The other is mounted inside the room; it measures the internal temperature. Both of the thermometers show the temperature analogue value.

Time is measured by using the wall clock. For most of the pilot cases the exact time is not quite important. Instead, students measure the amount of elapsed time. Figure 6 is a snapshot of the internal thermometer at 25°C and the clock at 6:09:49.

3.2 Educational Aspects

The design of the virtual classroom is based on a few iconoclastic decisions contributing substantially to successful pilot experiments:

- Continuous simulation;
- Dual purpose devices;
- Navigational freedom;
- Lack of internal description.

One of the most risky decisions is the implement continuous simulation. It starts immediately after the program is launched. There are no options to start, pause or stop the process of simulation.

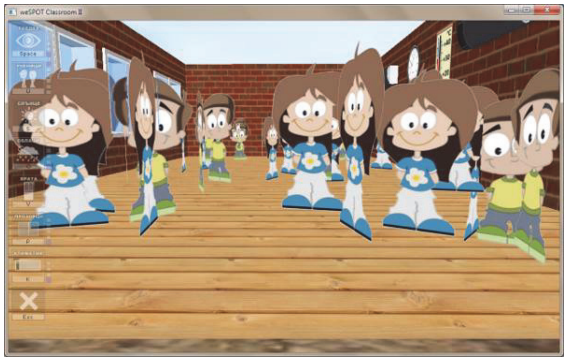


Figure 5: Overcrowded classroom.

The side effect of this decision is that temperatures and time are changed during the configuration of the scenario. The rationale for this decision is to provide real-life experience. This also forces the students to both accept some level of inaccuracy, as well as to think ahead and act promptly when they set up their experiments.

To provide even further feeling of real-life experiments the measuring devices (the clock and both thermometers) show their values in an analogue manner. There is no option to get the temperature or the current time digitally. Students have to "convert" analogue measures into numbers in real-time.

Additionally, the devices are not only output devices, i.e. they display measured values, but are also input devices – students are able to interactively change the internal/external temperatures and the current time. This decision aims at simplifying the interface and increasing the functionality.

While using the virtual classroom the students can navigate to any place around or inside the room. This design decision provides a feel of freedom. Students often spend some time just walking within the virtual scene and investigating every place. There is no documentation describing what exactly to do, where to go, how to measure, what is the impact of the weather, etc. Students are expected to explore not only a specific science problem, but also to explore the virtual world and the software.

4 A VIRTUAL CLASSROOM USED IN A REAL CLASSROOM

The experiments with the virtual classroom were held again with three classes of 6th graders in First Private Mathematical School in Sofia, Bulgaria. The weSPOT pilot itself started a week earlier, when the teacher in Human and Nature defined the research question to the students.

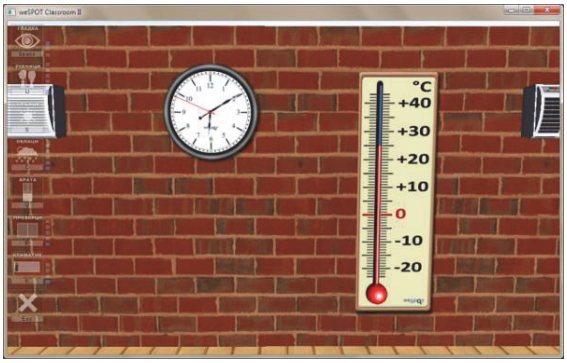


Figure 6: The internal thermometer and the wall clock.

The introduction of the Virtual Classroom simulation to the students was observed and

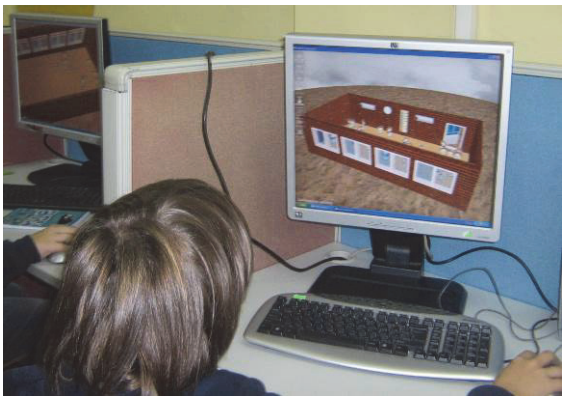


Figure 7: A student proving hypothesis.

documented by the Sofia University researchers, taking care for the Bulgarian pilots and providing support to teachers and students. The weSPOT pilots continue and their final results will be presented in next papers, but first observations and conclusions are presented below.

4.1 First Reactions

The Virtual Classroom was introduced to the students by teachers in Information Technologies, who are also members of the weSPOT school team. They presented the simulation only with several sentences, introducing it as a tool, which will support the students in their inquiry project for the Human and Nature discipline. In addition, teachers shared with students the fact, that the simulation is specially developed for their needs and purposes of the given inquiry. Then the students were left to discover the software functionalities.

Immediately each student started own inquiry with great power of observation in many different directions. Their reactions like: "*Great graphics*!", "*A-a-a, there is sun also!*", *"Oh! What students!*", "*The temperature drops very quickly when windows are widely open! Let's open them in the other direction.*" confirm this.

Each of the inquiries provoked the students' curiosity - they asked questions and defined their own hypothesis.

4.2 Open Questions and Hypotheses

Some of the questions like "*Why the room is without roof? Would this cause the outside and inside temperatures to mix?*" shows that the students expect the simulation to be completely realistic. Other questions like "*Why the classroom is on Mars*?" provided researchers with new ideas: to think about future use of Virtual Classroom for simulation of conditions at other planets.

However most of the questions were really dedicated to the inquiry, which students have worked on. In one of the cases students opened the door and all the windows of the Virtual Classroom, set the outside temperature at minimum and made the hypothesis that inside temperature will fall down quickly. But according to the simulation, the inside temperature has even slightly increased. "*Why?*" asked a student. Then the teacher invoked him to look carefully what factor could influence the temperature. "*Oh!*" was next reaction. "*In the room there are 40 pupils! Let's them go out and then check the temperature again.*" Through several experiments this student proved newly appeared hypothesis (Figure 7).

In some of the cases students had seemingly opposite hypotheses: If the door and the windows are open then: a) the room gets colder (Figure 8, top); b) the room gets warmer (Figure 8, bottom).

Two teams of students tried to argue who is right and why. Then they discovered that it is one and the same hypothesis, because in the first case the outside temperature was lower than the inside one; in the second case the outside temperature was higher than the inside one.

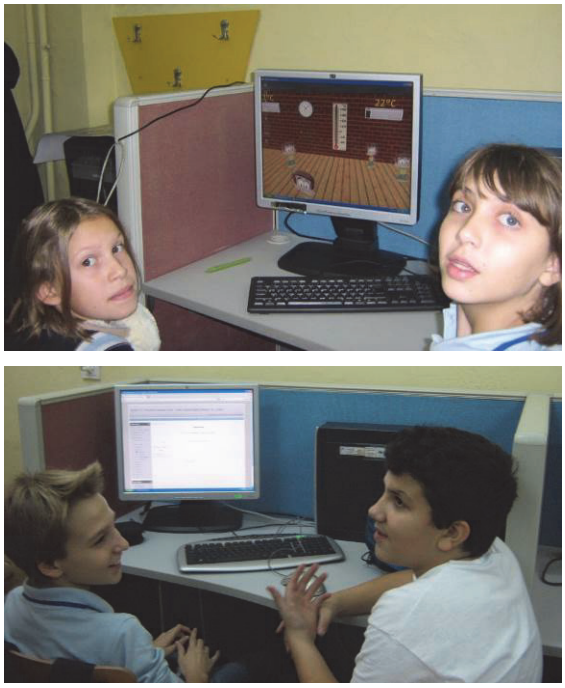


Figure 8: Students with first hypothesis (top photo) and students with "opposite" hypothesis (bottom photo).

4.3 Observations

The simulation provoked students' interest. Most of them expressed their expectation to run the program at home. The Virtual Classroom created atmosphere in which all the students worked intensively, but each one of them in own direction, according personal perception of the world, with own pace.

The students were excited by the simulation functionality, especially by the possibility to manipulate objects and phenomena. They gradually discovered, in different order, functionality of the Virtual Classroom elements: the status of the buttons, the change of the temperature, the rotation of the building with the Virtual classroom, going into the classroom and outside it, changing the number of pupils inside, etc.

The initial enthusiasm established better conditions for successful inquiry learning activities. After that, the students started collection of data through subsequent experiments, systematically proving their hypothesis and recording the data and conclusions in Moodle, where the special course for the purpose of the pilot was running.

5 CONCLUSIONS

The analysis of the results of the second pilot experiment shows that the Virtual Classroom simulation plays successfully its role in the weSPOT project. To what extend the software simulation contributes to the core science experiment objectives will be possible to conclude later, after the third pilot experiment, but even at this stage it was proven that students learn some of the important lessons like:

- If you have hypothesis, through collection of data you could prove or reject it.
- If you would like to confirm that given result is not accidental, then you need to receive the same result with the same data when recreating the same experiment.

During the next year the weSPOT inquiry workflow engine will be fully developed, providing opportunity to conduct inquiry, fully supported by technology, collecting and sharing in electronic form data and evidences. The third pilot will use the inquiry workflow engine together with the simulation and specially developed validating instruments. In this way we will be able to implement the whole inquiry cycle and application of the weSPOT model for IBL, and to evaluate its real contribution to learning.

ACKNOWLEDGEMENTS

The research is done with financial support of EC FP7 /2007-2013/ contract N°318499 – project weSPOT.

REFERENCES

Brown, J., Adler, R., 2008. Minds on Fire: Open Education, the Long Tail, and Learning 2.0, *EDUCAUSE Review, vol. 43*, pp. 16-32.

Kolb, D., 1984. *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., Protopsaltis, A., Held, P., Hetzner, S., 2012. weSPOT: A Cloud-based Approach for Personal and Social Inquiry, *WCLOUD 2012 Workshop on Cloud Education Environments*, editors: Alexander Mikroyannidis, Rocael Hernandez Rizzardini, Hans-Christian Schmitz, published in:Sun SITE Central Europe, 2012, pp. 7-11, Ref: http://ceur-ws.org/Vol-945/ws.org/Vol-945/

Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., Boytchev, P., Protopsaltis, A., Held, P., Hetzner, S., Kikis-Papadakis, K., Chaimala, F., 2013, weSPOT: A Personal and Social Approach to Inquiry-Based Learning, *Journal of Universal Computer Science (JUCS),* vol. 19, no. 14 (2013), pp. 2093-2111.

Mulholland, P., S. Anastopoulou, T. Collins, M. Feisst, M. Gaved, L. Kerawalla, M. Paxton, E. Scanlon, M. Sharples, and M. Wright, 2012. nQuire: technological support for personal inquiry learning. *IEEE Transactions on Learning Technologies*, 5 (2), pp. 157-169.

Stamenov, Sv., Dimitrova, T., 2013. My classroom – Energy Effective: Inquiry-based Science Education in the Classroom, *EDULEARN13 Proceedings of 5th International Conference on Education and New Learning Technologies*, 1-3 July, 2013, Barcelona, Spain, Edited by L. Gomez Chova, A. Lopez Martinez, I. Candel Torres, International Association of Technology, Education and Development, Published by International Association of Technology, Education and Development (IATED), Barcelona, Spain, pp. 5475-5483.

weSPOT, 2012. Working Environment with Social and Personal Open Tools for inquiry based learning, Research project funded from EC under FP7, 2012-2015, contract number 318499, http://portal.ou.nl /web/wespot.

STEM Education
Students Touch Space Through Free Robotics Programming Competition

Jenny Liu[1], Wendy Feenstra[2], Alvar Saenz Otero[1], and Kathleen Magrane[3]
[1]*Space Systems Laboratory, Massachusetts Institute of Technology,*
77 Massachusetts Avenue, Cambridge, MA, 02139, U.S.A.
[2]*Aurora Flight Sciences, 4 Cambridge Center Suite 11, Cambridge, MA, 02142, U.S.A.*
[3]*Massachusetts Afterschool Partnership, 128 Tremont Street-Suite 4F, Boston, MA, 02108, U.S.A.*
{liuj, alvarso}@mit.edu, wfeenstra@aurora.aero, kmagrane@massafterschool.org

Keywords: Middle School, Programming, Competition, Robotics, International Space Station, SPHERES, NASA.

Abstract: This position paper discusses the use of the Zero Robotics Summer Program competition as a tool to teach programing to students in middle school (ages 11-14). The benefits of teaching programing at that age are discussed, including grasping the student attention early on and providing a useful skill. Zero Robotics is a free programming competition where students program microsatellites to compete with other student-programmed satellites in a virtual game using an online simulation. The finalists see their code compete on the SPHERES satellites aboard the International Space Station, during a final competition transmitted live. The game and curriculum teaches students physics, mathematics, and programming concepts. This position paper argues that it is possible to engage students in programing by providing the right incentive (space) through a competitive environment where teachers have a curriculum to back them up, without becoming a standard classroom exercise.

1 INTRODUCTION

Zero Robotics Middle School Program is a programming competition which culminates with the top teams demonstrating their software aboard the International Space Station (ISS). The Synchronized Position Hold, Engage, Reorient, Experimental Satellites (SPHERES) are used inside the ISS, controlled by programs developed by grade-school students. The SPHERES satellites have operated aboard the ISS since 2006 to carry a wide range of algorithm development by graduate students (Saenz-Otero, 2005) (Wertz, 2000). Starting in 2009 the Zero Robotics program has leveraged the resources of SPHERES as a tool for Science, Technology, Engineering, and Math (STEM) education through a unique student robotics competition (Saenz-Otero, 2010).

Participants in Zero Robotics write their programs using an online programming environment accessed via a web browser, which does not require the installation of any special tools, nor the purchase of software licenses or entry fees. Thus, the "zero" stands for "zero-cost" and "zero-configuration" in addition to "Zero"-gravity. This no-cost model eliminates potential barriers that often limit the participation of under-served and under-represented populations in robotics competitions.

The program has grown over four years into an international tournament involving more than two thousand student competitors and has given hundreds of students the experience of running experiments aboard the ISS. The main goal of the program is to teach students STEM topics by showing them real world applications and opportunities for future involvement, to increase their awareness for space and technology, and to encourage them to explore these topics.

National Aeronautics & Space Administration (NASA), Defense Advanced Research Projects Agency (DARPA), and Center for the Advancement of Science in Space (CASIS) sponsor the Zero Robotics Middle School Summer Program. It is run by the MIT Space Systems Laboratory (SSL), the Massachusetts Afterschool Partnership (MAP), Aurora Flight Sciences Corporation (AFS), and Top Coder (TC). The Education Development Center, Inc. (EDC) led the 2013 curriculum development.

In 2013 the Zero Robotics Middle School Program expanded outside the state of

Massachusetts to include California, Florida, Georgia and Idaho. The 2010 Middle School pilot program utilized MIT undergraduate mentors who became daily instructors at each of the ten pilot locations. As ZR expands, the MIT mentor model is not sustainable and new methods of developing program quality must be applied. The MIT SSL and MAP partnered with EDC to develop and pilot a downloadable and free Zero Robotics curriculum geared towards middle school summer programs. The guide provides a structured outline of objectives and activities to support summer educators in presenting lessons and key concepts.

This paper presents the results observed from high school competitions and how the lessons learned have been applied to the middle school program. This position paper argues that it is possible for middle-school students to get interested in programing if (a) an exciting incentive, such as space, is used as initial attraction (b) the students are presented with a challenge/competition, and (c) the teachers are provided with a curriculum that does not require them to be an expert in computer science.

2 BACKGROUND

There are two different types of programming competitions, short-term competitions and long-term competitions. Short-term competitions usually take place over a day or weekend, and long-term competitions take place over several weeks or months. In this position paper, we will focus on long term programming competitions. For the purpose of this paper, our definition of middle school is grades six through nine (ages 11-14).

Programming competitions targeted towards middle school students are important for the following reasons:

1) *Students Need to Learn Programming because Programming Opens up a lot of Doors for Future Opportunities.*

According to Code.org, there are plenty of opportunities for students that learn programming. They estimate that there will be 1.4 million computing jobs by 2020, which equates to about $500 billion opportunity.

2) *Programming should be taught Starting in Middle School.*

Students must get involved in engineering and programing at a young age. There is precedent globally that students as young as first grade are capable of learning programming. Last year, the Estonian Tiger Leap Foundation (Wilson, R, 2012) launched a program called “ProgeTiiger” where programming is introduced to students starting in first grade. Estonia Tiger Leap Foundation decided to start this project because they saw how many companies struggle to find decent programmers. Estonia is proud of its programmers developing Skype, which has been acquired by Microsoft. ProgeTiiger starts in first grade. Therefore, it is not too early to start in middle school.

3) *We Believe that Teaching Students to Program by Competition is an Effective Method.*

The Zero Robotics Middle School Competition is based on the idea of gaming and wining. At the same time, it includes multiple features that encourage *friendly* competition. Field trip events bring students together, introducing students with similar interests to each other, creating a community of student programmers. Within this community, students support each other through messages in the Zero Robotics forum. This gives students a glimpse of collaboration, highly relevant to future careers at universities and in industry. In addition, the long-term competition format, gives students the opportunity to create and evolve their solutions similar to processes in universities and in industry. Microsoft Education blogger notes that long term programming competitions make students design well thought-out solutions that closely resemble industry or formal research tool development. (Microsoft, 2013)

2.1 Thesis

It is possible for middle-school students to get interested in programing if (a) an exciting incentive, such as space, is used as initial attraction (b) the students are presented with a challenge/competition, and (c) the teachers are provided with a curriculum that does not require them to be an expert in computer science.

3 IMPACT OF HIGH SCHOOL ZERO ROBOTICS COMPETITION

The Zero Robotics High School competition has attracted over 1500 students since 2009. Its initial growth in teams showed that even with a small number of schools that teach programing in the USA (under 10% of public schools http://code.org/stats), there is an attractiveness in having space as an

incentive. Feedback from participants presents individual data points that demonstrate the impact of space:

"A teacher in a local middle school has followed the progress of team y0b0tics! from their early days as a middle school FLL [First LEGO League] team through their current participation in ZR, on a poster that hangs in his science classroom. The poster's theme is "Start here…" (middle school FLL) and "Go here…" competing in ZR with your code being run by an astronaut on the ISS. Recently, my older son, Richard, attended an event at my younger son's school. While walking through the hall, a middle schooler approached him and said, "[You're] cool!" and explained recognizing him from the poster.

Mt. Olive High School in New Jersey, boasts a football team with 100+ participants and a robotics team with 125+ participants."

Other feedback addresses the ability to introduce new STEM concepts to students through the competition: "I never thought about programming, and wasn't interested in programming at all. However, the thought that I could control robots in space, especially in the International Space Station, caught my interest. And so I participated in Zero Robotics." The space incentive of ISS encourages and attracts students who never thought they could program or would be interested in learning programming. The program growth supports this positive impact. Zero Robotics started with only two teams in 2009, a total of thirteen students, and has expanded this year to 142 teams (includes both USA and European teams) with more than 1600 students.

The competition serves as an effective and incentivizing educational tool when compared to the traditional curriculum: "Several parents and students have expressed to me that ZR competition was one of the most important educational experiences of their entire career. In my opinion, students who participate have far greater understanding of physics, math, and programming than their peers after just a few weeks of a competition…I saw a big improvement in one student who joined the Zero team. For one thing, she now really understands vector concepts far better than other students, and vectors are fundamental to many different concepts in physics so this directly lead to her grades improving. It's on thing to memorize how to take a vector sum or a cross product. [It's] something else entirely to have to use those concepts to solve real problems you are faced with."

In addition, the competition format creates incentives for the students to dig deeper into the subjects: "The competition atmosphere encouraged me to learn as much as I could. Instead of just learning at team meetings, I looked up exercises and tutorials on my own time. Consequently, after the Zero Robotics season was over, I decided to take a formal computer science course. I felt that taking a formal course would enforce and expand upon concepts I gathered through on-the-fly coding during the Zero Robotics season.

"From the first day of my computer science class, I instantly noticed a distinct difference between coding concepts learned in school versus the ones I picked up over the course of Zero Robotics. The biggest difference would definitely be the motivation behind learning in the two different formats (Zero Robotics and traditional classroom setting). I was much more motivated in Zero Robotics because I had to make sure I kept up with students from other teams. The latest Zero Robotics Challenge always posed an interesting problem that required a solution that needed to be innovative. Exercises given in class were often tedious, lacking any real application to tasks one would do in the real world. As a result, I found the classroom methods to be extremely boring. For people driven by problem-solving and creative approaches to learning, I would recommend giving Zero Robotics a try. It definitely provides a fun way to learn something completely new like programming."

By working in teams, rather than individual competitions, ZR enhances the learning experience by encouraging collaboration: "I learned C entirely through Zero Robotics. … I learned to code for ZR while learning to code in Java, which helped in that there is a lot of cross-over. Aside from the generalities, the most crucial thing I learned from programming with ZR is that there are a lot of ways to do a task, so find the most efficient one. … and in general write code that other people on your team have to be able to understand and add to was definitely one of the biggest things I took away from ZR. …about collaboration. Learning to code with others can be a big challenge because most people learn to code on their own, and everyone has their own coding style. Like I said before, there are many ways to write the same program. Everyone wants to write it their own way, and programmers in general are very protective of their code. Learning to accept others' ideas and ways of writing is a valuable asset. In this way, ZR helped me a lot with group projects in robotics club and team programming competitions…"

The successes of the High School Tournament were implemented in the Middle School Summer Program:

- Maintain space as an incentive, including participation of the finalists aboard ISS.
- Make programming the central STEM objective.
- Maintain a *team* competition format which encourages collaboration.

The middle school program differs from the high school program in that it takes place over the summer (five weeks) and provides more support to middle school teachers via a curriculum of structured weekly activities.

4 COUNTERARGUMENTS TO MIDDLE SCHOOL

Counterarguments exists against teaching programing at the middle school level and using competition as a teaching tool. Literature review of the topic yields the following arguments:

1) *Middle School Students are too young to Understand Programming.*

In online blogs, parents advocate that middle school students should "enjoy their childhood years" and not be tied to the computer screen. Students should learn to follow their instincts, use their imagination, and not be taught the rigid world of engineering design (Farr, 2013). There is also suggestion that learning programming languages may sacrifice a student's ability to learn foreign languages (Koerner, 2013).

2) *Students are too young to be Competitive.*

Teachers involved in the Zero Robotics program provided feedback for the *elementary* school level (ages 6-11) that the focus of our program should be on problem solving and not competition.

"These students worked hard to get their programs to work…it would be nice to just 'see' how they did.

"The competition wasn't as important to the students as the sense of accomplishment from seeing that the student code worked."

4.1 Refutes to Counterclaims

Current events and industry activity review refutes the counterclaims:

- 17-year-old British developer Nick D'Aloisio's startup was acquired by Yahoo! D'Aloisio taught himself to code at age 12 (7th grade) and started developing iPhone apps in 2008. After launching his app, D'Aloisio received widespread attention from media and investors including actor Ashton Kutcher and Zynga founder Marc Pincus. (Fell, 2013)
- Microsoft held, Kodu, a game design programming competition, this year for students age nine through twelve. Since the competition, there have been one-half million downloads for the 16,000-plus kid-created games. (Microsoft, 2013)

5 ARGUMENTS IN SUPPORT OF MIDDLE SCHOOL

1) *Programming is an Important Skill for Middle School Students to Learn*.

As pointed out earlier, programming jobs are growing and there are many opportunities for students in the future. Thus, we believe that programming is an important skill for middle school students to learn. Further, by creating 'applications' for programing, students are exposed to other STEM skills such as math (algebra, geometry) and physics.

2) *Middle School Students are Developing Interests*.

Middle school students start developing hobbies before high school, therefore it is important to expose students to programming earlier. STEM research points out that one of the keys to successful education programs for students is to spark the interest of students in earlier grades. (Koerner, 2013)

Children are exposed to technology as early as age two. Toddlers play games and watch videos of themselves on iPads. Through these experiences, children are taught how to use them, but not how to program them. Not knowing how digital environments are constructed leads children to accept them at face value. Facebook is an example where kids think the function of Facebook is to help them keep in touch with friends. In actuality, Facebook users are now actually the product, the target of ads and marketing. Students should learn how to make tomorrow's software, not just about how to use today's software. (Rushkoff, 2012)

3) *Afterschool Programming Offers Opportunities for Middle School Students*.

Programming seems daunting to many students. It is often perceived as a hobby that only 'geeks' and 'nerds' pursue. Generally, students don't realize they are capable of learning to program themselves and shun away from it. Only a small percentage of students are exposed to computer science, and only 23% of those are minority. (http://code.org/stats)

STEM Afterschool Programs focus on providing services to underrepresented populations in STEM

fields, and many also focus on providing girls with exposure to science and female role models. (Afterschool Alliance, 2011)

Feedback from participating teachers reflects these statistics: "there is clearly a gender gap in these activities. That said, I believe that the more competitions/games that encompass STEM activities students are exposed to at an early age will tend to reduce the gap."

Therefore Zero Robotics aims to make the competition as accessible to as much of the population as possible. The competition has no entry free, requires no setup, and is free in the sense that all you need is a computer that has a web browser to compete. The rare opportunity for students to have their code run by astronauts in space is geared specifically to minority and underrepresented students at the middle school level. The program attempts to show that anyone can learn to code by providing a rich set of free step-by-step tutorials on our website.

6 CONCLUSIONS

Through review of current events, literature research, analysis of ZR program growth and participation, and substantial feedback from participants, it is possible to conclude that:

- Space and astronauts provide a clear incentive for middle-school students to participate in a programing competition, even if they had never thought of programing as something they could or wanted to do.
- Using a competition format encourages not only participation but also collaboration.
- Students learn to relate to each other in new ways, by working together in a common goal using their minds.
- A curriculum with clearly prescribed activities allows middle-school teachers to guide their students in learning programing

Based on the success, the Zero Robotics Middle School Summer Program will continue to expand.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Kopelow, Wei Low, Terri Tomassi, Stephen Pendergast, and Neil Shankar for their dedication to the program, and for sharing with us their perspective on the Zero Robotics Program.

REFERENCES

Afterschool Alliance, et al, 2011, "STEM Learning in Afterschool: An Analysis of Impact and Outcomes", www.afterschoolalliance.org.

Farr, C, 2013, "What's the right age for kids to learn to code?", http://venturebeat.com/2013/06/25/whats-the-right-age-for-kids-to-learn-to-code-debate/

Fell, J, 2013, "Yahoo Acquires Teen Entrepreneur's News Gathering App, Summly", http://www.entrepreneur.com/article/226196.

Koerner, B., 2013, "Forget Foreign Languages and Music. Teach Our Kids to Code.", http://www.wired.com/opinion/2013/09/ap_code/

Microsoft, 2013, "Microsoft Introduces Imagine Cup Competition for Young Students", http://www.microsoft.com/en-us/news/press/2013/mar13/03-19koduchallengepr.aspx.

Reville, P, 2010, "Closing the Inspiration Gap: A Major STEM Challenge", www.mass.gov.

Rushkoff, D, 2012, "Teach U.S. kids to write computer code", http://www.cnn.com/2012/12/10/opinion/rushkoff-code-writing/index.html?hpt=hp_bn7.

Saenz-Otero, A, 2005, "Design Principles for the Development of Space Technology Maturation Laboratories Aboard the International Space Station", *Doctoral Thesis*, MIT, Cambridge MA.

Saenz-Otero, A, et al, 2010, "ZERO-Robotics: A Student Competition aboard the International Space Station", *IEEE Aerospace Conference.*

Thomasian, J, 2012, "The Role of Informal Science in the State Education Agenda", www.nga.org.

Thompson, A, 2010, "Programming Contests – for good or for ill", http://blogs.msdn.com/b/alfredth/archive/2010/02/11/programming-contests-for-good-or-for-ill.aspx.

Wertz, J, Chen, A, et al, 2000, "SPHERES: A Testbed for Long Duration Satellite Formation Flying in Micro-Gravity Conditions", *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting*, AAS 00-110, Clearwater, FL.

Wilson, R, 2012 "Computer programming will soon reach all Estonian schoolchildren", http://ubuntulife.net/computer-programming-for-all-estonian-schoolchildren/

An Alternate Learning Approach for Destructive Testing in Civil Engineering

Benefits from Remote Laboratory Experimentation

Xuemei Liu, Nannan Zong and Manicka Dhanasekar
School of Civil Engineering and Built Environment, Queensland University of Technology,
2 George St, Brisbane, QLD, 4001, Australia
{x51.liu, roger.zong, m.dhanasekar}@qut.edu.au

Keywords: Alternative Learning Approach, Destructive Testing, Civil Engineering Education, Remote Laboratory Experimentation.

Abstract: An alternative learning approach for destructive testing of structural specimens in civil engineering is explored by using a remote laboratory experimentation method. The remote laboratory approach focuses on overcoming the constraints in the hands-on experimentation without compromising the understanding of the students on the concepts and mechanics of reinforced concrete structures. The goal of this study is to evaluate whether or not the remote laboratory experimentation approach can become a standard in civil engineering teaching. The teaching activity using remote-laboratory experimentation is presented here and the outcomes of this activity are outlined. The experience and feedback gathered from this study are used to improve the remote-laboratory experimentation approach in future years to other aspects of civil engineering where destructive testing is essential.

1 INTRODUCTION

Civil engineering is a practical discipline which applies scientific and mathematical principles in a socially responsible manner to design, construct, and operate infrastructures and building systems. The overall goal of the civil engineering education is to prepare students for the profession to get solutions for the practical problems. To do this successfully, civil engineers must have the knowledge that is traditionally gained in the educational laboratories (Feisel and Rosa, 2005).

Laboratory based courses play a critical role in engineering education. Nersessian (1991) claims that "hands-on experience is at the heart of science learning" and Clough (2002) declares that laboratory experiences "make science come alive". Lab courses have a strong impact on students' learning outcomes, according to Magin et al., (1986). Instructional laboratories have been implemented for engineering education from the early days. The traditional one is known as hands-on laboratory with real instruments. Feisel and Rosa (2005) summarised the fundamental objectives of the engineering teaching laboratories, which can be used as a guide for engineering educators to develop and improve the effectiveness of laboratory learning experiences. There are three types of educational laboratories in engineering education (Ma and Nickerson, 2006). These include hands-on laboratory, simulated or virtual laboratory, and remote or distributed learning laboratory (Krivickas and Krivickas, 2007).

Remote/ virtual laboratories are currently being developed and used in many places around the world in areas that do not require destructive testing. There have been debates over the introduction of remote/ virtual versus hands-on laboratories in engineering education. Hands-on laboratory allows students to directly see, hear, touch, and feel the devices and the experimental specimens. Whilst in the virtual laboratory students learn engineering principles through simulation running on computers without any real element of equipment or specimens. In the remote laboratory, students interact with the real devices/ equipment/ specimens remotely through a computer user interface. Therefore, the remote laboratories are called as the "Second Best of Being There" by Aktan (1996). However, with the rapid advancement in technologies, even hands-on laboratory utilises more and more computers and technical devices and controllers which blurs the

hands-on and the remote laboratory learning experiences.

There is still a shortage of data on whether new technologies such as remote laboratory experimentation are as effective as hands-on laboratory when it comes to teaching design skills involving destructive testing of materials/ structural members. The effectiveness of the remote-laboratory compared with traditional hands-on laboratory practice is seldom explored. Therefore, this paper will discuss the effectiveness of the remote laboratory experimentation for civil engineering undergraduates through an analysis of the students' feedback based on a remote-laboratory project. The main target is to promote an effective use of alternative learning approach for undergraduate learning civil engineering.

2 REMOTE LABORATORY EXPERIMENTATION

An year 2 undergraduate subject, ENB 276 Structural Engineering-I, that aims at introducing the analysis of simple statically indeterminate structures, pattern loadings in structural design and the behaviour and design of reinforced concrete beams, slabs, and columns is used for the introduction of remote lab testing. Laboratory practice is an important element for this unit. Historically hands-on laboratory practices were implemented for students to design and construct reinforced concrete beams and test them to failure within the on-campus laboratory. In those days, the student number was around 90. With the relocation of the on-campus laboratory to a new campus in a remote suburb, the hands-on laboratory practices become less efficient for a cohort with a large number of students (398 enrolments). In order to enable the students to experience what they would do for the hands-on experiment in the laboratory, a remote laboratory project was developed as explained in this paper.

2.1 Background

The idea was that the students design their own beams in a team of 4-5 members of their choice. The criteria (capacity of the testing machine and space allowance) of the design were explicitly stated in the design brief. Basically, one beam was designed for bending failure and the other for shear failure. Two beams among all the designed beams were then selected as the test specimens for the remote laboratory experimentation.

The reinforced concrete beams were prepared by tutors and technicians. The whole preparation process from formwork preparation, steel bar bending, placement and positioning, concrete materials proportioning, mixing, and lastly the casting of concrete beams were all video recorded and played in the lecture theatre before the remote laboratory testing. At the time of testing, the students sitting inside the lecture room could remotely control the testing machine through an internet protocol (IP) and observe the whole testing procedure through the live streaming of an IP camera, while technicians in the laboratory are supervising the whole testing process in case of hazard events happening and having real-time communication with the students remotely through another camera. The detailed test setup is narrated in the following section.

At the end of the semester, after the final examination and declaration of result, the students were surveyed on different aspects of the influence of the remote laboratory experimentation on their learning experience and outcomes through a voluntary online questionnaire system. The feedbacks based on a respondent number of 53 (out of 398, or 13.3%) are used in the analysis of the effectiveness in the learning experience and learning outcomes by using remote laboratory experimentation in civil engineering education. The low response rate is typical at Queensland University of Technology as the students are surveyed (response is voluntary) for each subject in each semester by a university wide system known as "Reframe". Further the questionnaire from this unit was personally carried out using blackboard system.

2.2 Remote Laboratory Setup

The overall remote laboratory setup is presented in Figure 1 schematically. The students sitting inside the lecture room remotely operated the controller of the actuator that applied loading on the test specimen in the laboratory. The remotely controled panel image was projected onto Projector 1 in the lecture theatre. Meanwhile, performed testing was captured by an IP camera, the live streaming images were used to feedback to the lecture theatre (PC3), and projected onto Projector 2. In addition, the mutual communication was established in parallel between the students in the lecture theatre and the staffs in the laboratory through the use of Skype for cost-effectiveness. This is to make sure the information from both sides can be instantly exchanged and the whole process can be conducted

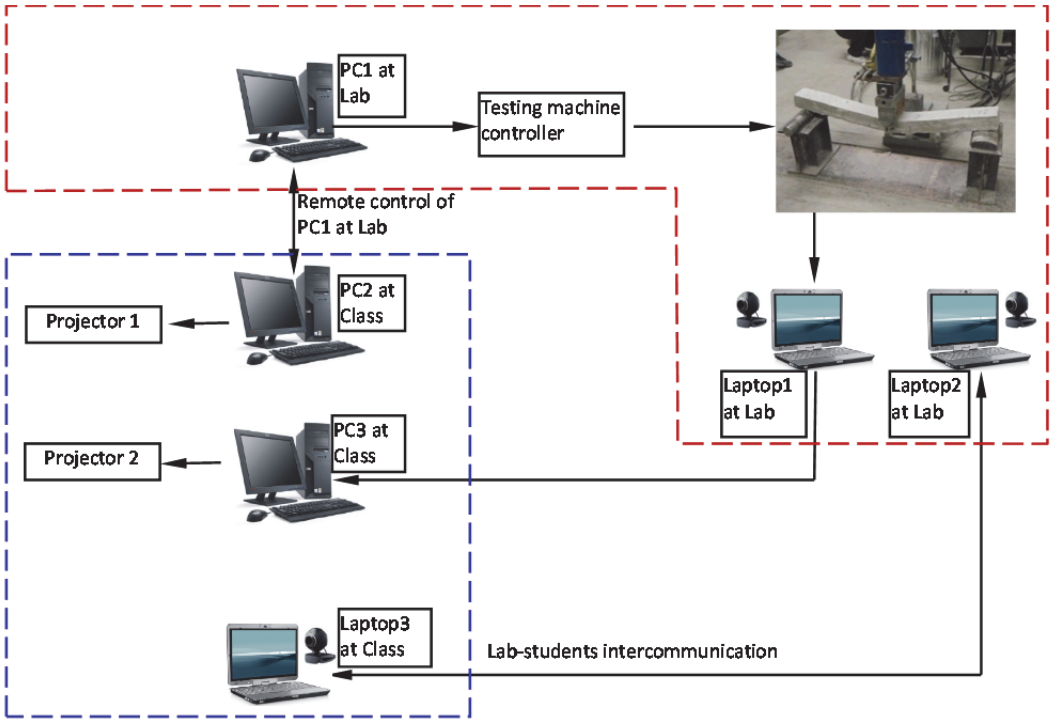


Figure 1: Schematic setup of the remote laboratory experimentation.

seamlessly and safely.

The lecture theatre was quite full on the day of remote testing with approximately 90% attendance and keen participation; in contrast during theory lectures attendance varied between just 60% - 75%.

2.3 The Benefits of Remote Laboratory

2.3.1 The Effectiveness of Video Casting in Learning Outcomes

The first part of the remote laboratory project involved the learning experience using webcast video. The video directly presents proportioning of materials for concrete, mixing and construction of beams. The benefits of the project in mix proportioning and mixing concrete by using the webcast video is shown in Table 1.

Table 1: The benefits of remote laboratory in mix proportioning and mixing concrete.

Benefits	Response
I know how to do mix proportioning	13%
I know how to mix concrete	4%
I know both	79%
None	4%

In the result, not surprisingly, the students benefitted a lot from the webcast video in extending their understanding of the mix proportioning and mixing of concrete. 79% of students responded that they have now known both mix proportioning and the mixing process of concrete. While 13% of them only knew how to do the mix proportioning and 4% of them only knew the mix procedure of concrete mixing. Although 4% of the students still admitted that they knew nothing about the mix proportioning and mix procedure of concrete, the benefits of the video cast are imminent for improvement of student learning outcomes.

2.3.2 The Outcomes of Remote Laboratory Learning

The benefits of designing reinforced concrete members as part of the remote laboratory experimentation for the students are summarised in Table 2 based on the feedback. It should be reminded that the remote laboratory only involved the design and testing of reinforced concrete beams. In the feedback, 88% of the students admitted that they have benefited from this project in designing the reinforced concrete beams in comparison to other reinforced concrete members (such as slabs) columns and slabs which were not covered in the remote laboratory project. This could therefore be inferred that the remote project greatly and directly has strengthend students' understanding and

impression of designing concrete beams. As other structural elements have not been tested, students found the behaviour of those elements more challenging. The results could be improved if the students not only design their specimen but also make the specimens by themselves, unfortunately it was difficult to involve students in preparing specimens due to time and space constraints. Negative data from 4% of them came with explicit statement of them not getting benefits in either designing or testing of any of the reinforced concrete members, namely, beams, slabs, or even columns as presented in Table 2.

Table 2: The benefits in designing reinforced concrete members.

Benefits	Response
Beams	88%
Slabs	6%
Columns	2%
None	4%

The two reinforced concrete beams with different reinforcement arrangement providing two different failure modes – bending and shear - were purposely planned to help students understand the differences between these two failure modes of reinforced concrete beams. 87% of the students strongly agreed that they have clearly understood these two different failure modes. About 12% students got some benefits in understanding of the failure modes of the reinforced concrete beams. While only 2% thought that they did not get any benefits regarding the failure modes of the reinforced concrete beams from the remote laboratory experimentation. As shear failure is more brittle, it is good to know a high majority of students could differentiate the two (flexure and shear) failure modes.

Table 3: The benefits in clearly understanding the different failure modes of the reinforced concrete beams.

Benefits	Response
Yes, absolutely.	87%
A bit	12%
Not at all	2%

2.3.3 The Benefits of Remote Laboratory in Comparison to Hands-on Experiment

Regarding to the general benefits from the remote laboratory experimentation to the students learning experiences, they gave different opinions. About 65% of the students agreed that the remote laboratory experimentation improved the learning outcomes of the engineering study. 50% of them thought that the remote laboratory experimentation will benefit them in implementing new technologies into their study and work. 44% of the students believed that this project benefited them in working as a team. About a quarter of them (29%) believed that this remote laboratory experimentation project improved their skills in organising reports. The results are shown in Table 4.

Table 4: The benefits of remote laboratory experimentation for engineering education.

Benefits	Response
Working as a team	44%
Organizing reports	29%
Implementing new technology into study and work	50%
Improving learning outcomes of engineering study	65%

3 CONCLUSIONS

The remote laboratory experimentation has not been often utilised for education in civil engineering design units involving destructive testing of material and structural specimens. Although the most desirable option is “hands-on” experimentation, with relocation of heavy structural labs away from city campus into suburbs and with very large cohorts, it becomes not possible to offer the hands-on approach; therefore, remote-lab is the most feasible option. There is no evidence of utilising this approach for experiments involving destruction of material and structural specimens. The information provided in the paper can therefore be considered as first of its kind for destructive testing of RC beams.

A case study for teaching structural engineering (Reinforced concrete design) that involved both video casting and remote laboratory experimentations is presented. The remote laboratory experimentation involved team-based design of reinforced concrete beams subjected to different failure modes, construction of the beams, and testing by using the remote-laboratory setup. The feedback on the understanding and the learning experience and learning outcomes are also presented. This feedback allowed to assess how well students benefitted and made use of the project, the video casting, and the remote laboratory experimentation.

It is concluded that the remote laboratory experimentation is an effective method for teaching and learning of subjects involving reinforced concrete design, where destruction of concrete cylinders and reinforced concrete beams are unavoidable. It creates an alternative learning

approach for the students by implementing new technologies. The use of webcast video and remote laboratory experimentation allows comprehensive learning of the structural engineering basics, construction of reinforced concrete beams and understanding of the failure modes (bending or shear). The effectiveness of the remote laboratory experimentation was confirmed by the students' positive feedback from the case study described in this paper.

In addition, the students' feedback will help us to shape the future teaching to further improve the teaching and learning experience of civil engineering education.

ACKNOWLEDGEMENTS

The authors acknowledge the support from Frank De Bruyne, technical manager of Banyo Pilot Plant Precinct (PPP), Queensland University of Technology. The assistance of several graduate students and technician, especially Sebastian Schundau and Shahid Nazir, is gratefully acknowledged.

REFERENCES

Aktan, B., Bohus, C. A., and Short, M. H., 1996. Distance learning applied to control engineering laboratories, *IEEE Transactions on Education*, 39(3).

Clough, M. P. 2002. Using the laboratory to enhance student learning. In *Learning Science and the Science of Learning*, Bybee, R. W., Ed. National Science Teachers Association, Washington, DC, 85–97.

Feisel, L. D., 2005, The role of the laboratory in undergraduate engineering education, *Journal of Engineering Education,* 94(1), 121-130.

Krivickas, R. V. and Krivickas, J. 2007. Laboratory Instruction in Engineering Education, *Global Journal of Engineering Education,* 11(2), 191-196

Magin, D. J., Churches, A. E., and reizes, J. A. 1986. Design and experimentation in undergraduate

mechanical engineering. In *Proceedings of a Conference on Teaching Engineering Designers*. Sydney, Australia Institution of Engineers, 96–100.

Nersessian, N. J. 1991. Conceptual change in science and in science education. In *History, Philosophy, and*

Science Teaching, Matthews, M. R., Ed. OISE Press, Toronto, Canada, 133–148.

How do Student Evaluations of Courses and of Instructors Relate?

Tamara Sliusarenko, Line H. Clemmensen and Bjarne Kjær Ersbøll
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet, Building 324, Kgs. Lyngby, Denmark
tasl@dtu.dk, lkhc@dtu.dk, bker@dtu.dk

Keywords: Course Evaluation, Teacher Evaluation, Canonical Correlation Analysis.

Abstract: Course evaluations are widely used by educational institutions to assess the quality of teaching. At the course evaluations, students are usually asked to rate different aspects of the course and of the teaching. We propose to apply canonical correlation analysis (CCA) in order to investigate the degree of association between how students evaluate the course and how students evaluate the teacher. Additionally it is possible to reveal the structure of this association. Student evaluations data is characterized by high correlations between the variables within each set of variables, therefore two modifications of the CCA method; regularized CCA and sparse CCA, together with classical CCA were applied to find the most interpretable model. Both methods give results with increased interpretability over traditional CCA on the present student evaluation data. The method shows robustness when evaluations over several years are examined.

1 INTRODUCTION

Teacher evaluations and overall course quality evaluations are widely used in higher education. Students usually submit their feedback about the teacher and the course anonymously at the end of the course. The results are usually employed to improve the courses for future students and to improve the instructor's effectiveness.

The research on student evaluations is important to make improvements in course construction and teaching methods. Student evaluation of teaching (SET) is a very well documented and studied tool. An overview of research on student ratings of instruction by Marsh (2007) demonstrates that student ratings are multidimensional, quite reliable, reasonably valid, and a useful tool for students, faculty and university administrators. The author also states that SETs primarily are a function of the instructor who teaches a course rather than the course that is taught.

Several studies on SET data investigate the relationship between student ratings and student achievements (Cohen, 1981; Feldman, 1989; Abrami et al., 1997). The main conclusion is that a student's achievement is correlated with a student's evaluation of the teacher and the course. Other issues, that are often discussed are relationships between the SET scores and various student-specific, course-specific and instructor-specific characteristics (Marsh, 1987).

This paper analyses the student evaluations from another angle; by investigating the correlation between how students evaluate the course and how students evaluate the teacher. The objective is to analyze the degree of association and in this way obtain a different angle on the perspective in Marsh's paper: that SETs primarily is a function of the instructor rather than the course. As a subject we have chosen to study a single course over time.

2 LITERATURE REVIEW

The most common method used to investigate the correlation amongst two sets of variables is canonical correlation analysis (CCA), introduced by Hotelling (1935). CCA can also be used to produce a model which relates the two sets of variables through linear combinations. The method has similarities with both multivariate regression and principal component analysis.

Application of CCA when variables in the sets are highly correlated or when the sample size is insufficient can lead to computational problems, inaccurate estimates of parameters or non-generalizable results. One way to deal with these problems is to introduce a regularization step into the calculations.

The first attempt to introduce the ridge regression technique, developed by Hoerl and Kennard (1970),

to the problem of canonical correlation analysis was proposed by Vinod (1976) and later developed by Leurgans et al. (1993).

The first development of Sparse CCA, a method that incorporates variable selection and produces linear combinations of small subsets of variables, was presented in Parkhomenko et al. (2007). They proposed an iterative algorithm that uses soft thresholding for feature selection. Waaijenborg et al. (2007) adapted the elastic net (Zou and Hastie, 2005) to canonical correlation analysis. Various approaches to introduce sparsity into the CCA framework were proposed in more recent works by LeCao et al. (2009), Witten and Tibshirani (2009), Hardoon and Shawe-Taylor (2011). Sparse CCA solves the problem of interpretability providing sparse sets of associated variables. These results are expected to be more robust and generalize better outside the particular study.

3 METHODOLOGY

3.1 Canonical Correlation Analysis

Canonical correlation analysis (CCA) was used to investigate the degree of association between the evaluation of the teacher and the evaluation of the course. CCA finds linear combinations of variables with the highest correlation between two sets of variables.

The method considers two matrices X and Y of order $n \times p$ and $n \times q$ respectively. The columns of X and Y correspond to variables and the rows correspond to experimental units. Classical CCA assumes $p \leq n$ and $q \leq n$. The main idea behind CCA is to find canonical variables in the form of two linear combinations:

$$\begin{aligned} w_1 &= a_{11}x_1 + a_{21}x_2 + ... + a_{p1}x_p \\ v_1 &= b_{11}y_1 + b_{21}y_2 + ... + b_{q1}y_q \end{aligned} \tag{1}$$

such that the coefficients a_{i1} and b_{i1} maximize the correlation between two canonical variables w_1, and v_1. In other words, the problem consists in solving

$$R_1 = corr(v_1, w_1) = \max_{a,b} corr\left(a^T X, b^T Y\right) \tag{2}$$

This maximal correlation between the two canonical variables v_1 and w_1 that are sometimes called canonical variates, is called the first canonical correlation. The coefficients of the linear combinations are called canonical coefficients or canonical weights.

The method continues by finding a second set of canonical variables, uncorrelated with the first pair that has maximal correlation. Wilks's lambda is used to test the significance of the canonical correlations.

Figure 1 illustrates the variable relationships in a hypothetical CCA. To answer the question "which variables are contributing to the relationship between the two sets?" the standardized canonical weights (i.e. coefficients used in linear equations that combine observed variables into latent canonical variable) and structure coefficients, also called canonical factor loadings, (i.e. correlations between observed variables and latent canonical variables) for the first significant canonical dimensions should be investigated.

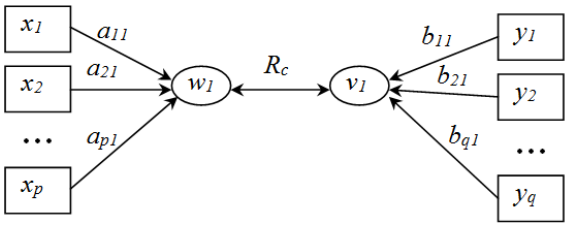


Figure 1: Visualization of CCA results.

Canonical correlation analysis helps to identify the major association between evaluation of the course and evaluations of the teacher. To perform classical CCA the R package CCA, developed by Dejean and Gonzalez (2009) was used. The package is freely available from the Comprehensive R Archive Network (CRAN) at www.r-project.org

3.2 Regularized Canonical Correlation Analysis

CCA cannot be performed when the variables $x_1, x_2, ..., x_p$ and/or $y_1, y_2, ..., y_q$ are highly correlated. In this case the correlation matrices, that are used in the computational process, tend to be ill-conditioned and their inverses unreliable. To deal with this problem a regularization step can be included in the calculations.

In CCA the regularization is achieved by adding a corresponding identity matrix multiplied by a regularization parameter to the correlation matrices.

$$\begin{aligned} \Sigma_{XX}(\lambda_2^a) &= \Sigma_{XX} + \lambda_2^a I_p \\ \Sigma_{YY}(\lambda_2^b) &= \Sigma_{YY} + \lambda_2^b I_q \end{aligned} \tag{3}$$

As the result the matrices become non-singular and a unique solution can be achieved. In order to choose "good" values of regularization parameters λ_2^a and λ_2^b , the k-fold cross-validation procedure can be used (González et al., 2009).

3.3 Sparse Canonical Correlation Analysis

Sparse CCA (SCCA) is an extension of CCA that performs a selection of variables jointly with the analysis of the two data sets. SCCA can also help to solve the problem of interpretability providing sparse sets of associated variables by setting the canonical correlation weights to zero.

As was mentioned above, CCA finds the vectors a and b, that maximizes $corr(a^T X, b^T Y)$. The way to obtain penalized canonical variates is to impose L_1 penalties on vectors a and b. So the optimization problem can be written as:

$$\begin{aligned} &\max_{a,b} \left(a^T X^T Y b\right) \\ &s.t. \|a\|_2^2 \leq 1, \|b\|_2^2 \leq 1, \|a\|_1 \leq \lambda_1^a, \|b\|_1 \leq \lambda_1^b \end{aligned} \tag{4}$$

This problem can be solved using the penalized matrix decomposition (PMA) approach, proposed by Witten et al. (2009). When λ_1^a and λ_1^b are small, some elements of a and b will be exactly zero. The algorithm yields sparse vectors a and b that maximize $cor(Xa, Yb)$. Values of regularization parameters λ_1^a and λ_1^b can be chosen using cross-validation.

To perform sparse CCA, the R package "PMA" (Witten et al., 2013) was used. The package also contains a function that helps to select tuning parameters by using cross validation.

4 DATA DESCRIPTION

Students at the Technical University of Denmark regularly evaluate courses by filling in web-forms a week before the final week of the course. The evaluations are intended to be a tool for quality assurance for: teachers, the department educational boards, and the department and university managements. On-line course evaluation at the university consists of three forms: Form A contains specific quantitative questions about the course, Form B contains specific quantitative questions about the teacher and Form C gives students an opportunity to write their qualitative feedback. This particular analysis is based on investigation of the relationship between answers in Form A and Form B.

- A.1.1 (Learning a lot): I think I am learning a lot in this course.
- A.1.2 (TM activates): I think the teaching method encourages my active participation.
- A.1.3 (Material): I think the teaching material is good.
- A.1.4 (Feedback): I think that throughout the course, the teacher has clearly communicated to me where I stand academically.
- A.1.5 (TA continuity): I think the teacher creates good continuity between the different teaching activities.
- A.1.6 (Workload): 5 points is equivalent to 9 hours per week. I think my performance during the course is.
- A.1.7 (Prerequisites): I think the course description's prerequisites are.
- A.1.8 (General): In general, I think this is a good course.
- B.1.1 (Good grasp): I think that the teacher gives me a good grasp of the academic content of the course.
- B.1.2 (Communication): I think the teacher is good at communicating the subject.
- B.1.3 (Motivate activity): I think the teacher motivates us to actively follow the class.
- B.2.1 (Instructions): I think that I generally understand what I am to do in our practical assignments/lab courses/group computation/group work/project work.
- B.2.2 (Understanding): I think the teacher is good at helping me understand the academic content.
- B.2.3 (Feedback): I think the teacher gives me useful feedback on my work.

The students rate the questions on a 5 point Likert scale (1932) from 5 to 1, where 5 corresponds to the student strongly agreeing with the underlying statement and 1 corresponds to the student strongly disagreeing with the statement. For questions A.1.6 and A.1.7, a 5 corresponds to too high and 1 to too low. In a sense for these two questions a 3 corresponds to satisfactory and anything else (higher or lower) corresponds to less satisfactory.

It is common practice to only include the first 3 questions for the teacher evaluation (B.1.1-B.1.3) for large courses. In such cases, the second part of the form B (questions B.2.1-B.2.3) is active for the teaching assistants only. Here, we examine one course "Introductory Programming with Matlab". The course is one of the largest courses at the university where all 6 questions from the teacher evaluation (Form B) are usually active.

The Introductory Programming with Matlab course is available 4 times per year: twice as a 13-week course (fall and spring semesters) and twice as an intensive 3-week course (January and June). The numbers of students that follow the course are very different from semester to semester. Here we will focus on the intensive 3-week version of the course. June courses are more popular (approximately 300

students) than the January courses (around 100-150 students). Figure 2 shows the number of students registered for the course and the course evaluation response rate over the period from January 2010 to June 2013.

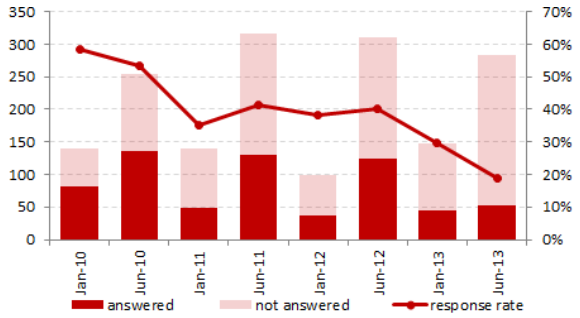


Figure 2: Number of course participants and evaluation response rate from January 2010 to June 2013.

For the comparison of methods we use results from one semester (January 2010), and for the robustness study we examine the same course at two other time points (June 2011 and June 2012). It should be noted, that students who participate in the course have very different backgrounds. Students at the university are obligated to take one programming course. Therefore the "Introductory Programming with Matlab" is a quite popular course among students on non-programming study-lines.

5 RESULTS

This section first presents a summary of the data, secondly the results of three versions of the canonical correlation analysis performed on the same data. Finally, the results of the robustness study are presented.

5.1 Evidence from the Data

The data set under investigation consists of 69 observations from the "Introductory Programming with Matlab" course held in January 2010. The course is one of the largest courses at the university, where the teacher is evaluated using all 6 questions from form B. Table 1 presents means and standard deviations of the answers from the responders. On average students gave rates below 3 to both the teacher and the course.

Question A.1.2 (I think the teaching method encourages my active participation) got the lowest average grade among course-related questions and question B.2.1 (I think that I generally understand what I am to do in our practical assignments) got the lowest average grade among teacher-related questions.

Table 1: Variable Mean and Standard Deviation.

Evaluation of the course			Evaluation of the teacher		
Question	Mean	St. Dev.	Question	Mean	St. Dev.
A.1.1	2.46	1.16	B.1.1	2.57	1.10
A.1.2	2.11	1.09	B.1.2	2.80	1.30
A.1.3	2.62	1.18	B.1.3	3.01	1.29
A.1.4	2.43	0.98	B.2.1	2.22	1.08
A.1.5	2.58	1.03	B.2.2	2.50	1.11
A.1.6	2.67	0.83	B.2.3	2.42	1.05
A.1.7	3.06	0.45			
A.1.8	2.36	1.06			

Table 2: Correlations among the Form A variables.

	A.1.1	A.1.2	A.1.3	A.1.4	A.1.5	A.1.6	A.1.7	A.1.8
A.1.1	1.00							
A.1.2	0.72	1.00						
A.1.3	0.57	0.48	1.00					
A.1.4	0.34	0.24	0.44	1.00				
A.1.5	0.55	0.54	0.53	0.45	1.00			
A.1.6	-0.34	-0.21	-0.11	0.07	-0.16	1.00		
A.1.7	0.01	0.19	0.21	0.24	0.15	0.29	1.00	
A.1.8	0.83	0.77	0.61	0.37	0.56	-0.24	0.08	1.00

Table 3: Correlations among the Form B variables.

	B.1.1	B.1.2	B.1.3	B.2.1	B.2.2	B.2.3
B.1.1	1.00					
B.1.2	0.81	1.00				
B.1.3	0.81	0.85	1.00			
B.2.1	0.47	0.49	0.43	1.00		
B.2.2	0.78	0.74	0.77	0.58	1.00	
B.2.3	0.64	0.57	0.67	0.55	0.78	1.00

Table 2 and Table 3 present the correlations between the variables from Form A and Form B. The correlations appear to be quite high especially within Form B. This can lead to uninterpretable results of classical CCA.

Figure 3 presents the average SET scores of the evaluation of the course (Form A) starting from January 2010 until June 2013.

There were some changes to the course during the period. In June 2010 the course was run by a new teacher, who introduced a new textbook, which seems much better than the Matlab notes used before as the course material got better feedback after this (A.1.3). Additionally, the course responsible team continuously works on improvement of the course and on making it less dependent on teacher and teaching assistants. Overall, there is a tendency of improvement of SET scores over the period from January 2010 to June 2013, with exception of the June 2012 semester, when the course got lower evaluation results.

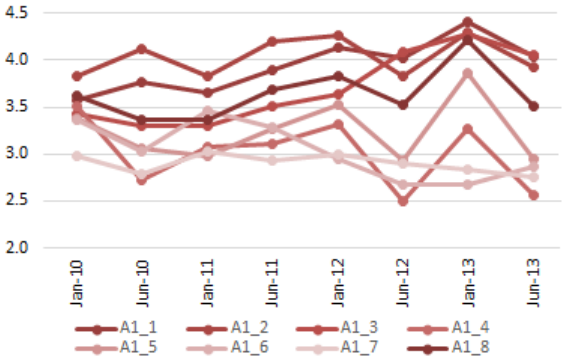


Figure 3: Results of the evaluation of the course from January 2010 to June 2013.

5.2 CCA Results

Figure 4 presents the canonical correlations and corresponding p-values for the significance test of each canonical correlation. In general the number of canonical correlations is equal to the number of variables in the smallest set. However, the test shows that only the first 4 canonical correlations are statistically significant. This means that the structure of the association between course and teacher evaluations lies in 4 dimensions, which is hard to interpret. The values of canonical correlations give an overall indication of a strong association between teacher and course evaluation.

Table 4 presents the standardized canonical coefficients and table 5 presents the correlations between the variables and their canonical variates. These coefficients are used to find the structure of the canonical correlations.

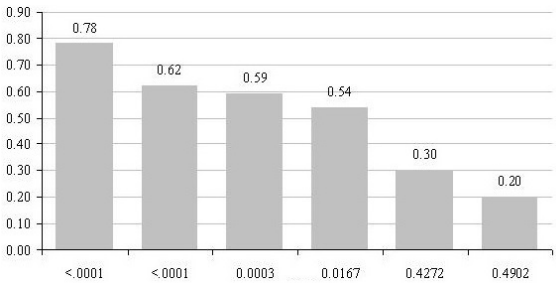


Figure 4: Canonical correlations and corresponding *p*-values.

For the first canonical correlation, questions A.1.5 (continuity between the different teaching activities) and A.1.8 (overall course quality) from the course related questions are the most important. Among the teacher related, question B.1.1 (good grasp of the academic content) is the most important. However, due to high correlations between questions within each set of variables, canonical factor loadings indicate that

Table 4: Standardized canonical coefficients.

Standardized Canonical Coefficients for the Form A variables					Standardized Canonical Coefficients for the Form B variables				
	V1	V2	V3	V4		W1	W2	W3	W4
A.1.1	-0.03	-0.31	0.86	-1.12	B.1.1	0.80	-0.86	0.65	-0.77
A.1.2	-0.16	0.08	0.33	0.34	B.1.2	0.28	1.37	-0.33	0.75
A.1.3	0.34	0.90	-0.12	-0.70	B.1.3	0.18	-0.25	-1.19	-0.28
A.1.4	-0.10	-0.50	0.78	0.06	B.2.1	-0.03	0.68	0.61	-0.08
A.1.5	0.60	-0.67	-0.54	0.26	B.2.2	-0.17	-0.25	0.59	-0.72
A.1.6	-0.11	0.09	0.35	0.08	B.2.3	-0.09	-0.48	0.06	1.55
A.1.7	-0.12	0.39	0.16	0.39					
A.1.8	0.42	0.38	-0.55	1.17					

Table 5: Canonical structure.

Correlations between the Form A variables and their canonical variables					Correlations between the Form B variables and their canonical variables				
	V1	V2	V3	V4		W1	W2	W3	W4
A.1.1	0.73	0.01	0.42	-0.16	B.1.1	0.97	-0.14	0.20	0.01
A.1.2	0.60	0.15	0.33	0.31	B.1.2	0.89	0.33	-0.06	0.20
A.1.3	0.76	0.49	0.24	-0.23	B.1.3	0.87	0.00	-0.21	0.21
A.1.4	0.38	-0.24	0.71	0.11	B.2.1	0.42	0.43	0.62	0.24
A.1.5	0.87	-0.28	0.05	0.19	B.2.2	0.71	-0.07	0.35	0.20
A.1.6	-0.35	0.17	0.32	0.25	B.2.3	0.56	-0.24	0.29	0.69
A.1.7	-0.02	0.43	0.37	0.48					
A.1.8	0.79	0.18	0.26	0.25					

the questions: A.1.1, A.1.2, A.1.3 from Form A and questions: B.1.2, B.1.3, B.2.2, B.2.3 from Form B are also important for the first canonical correlation. The structures of the other 3 canonical correlations can be found by similar analyses of the corresponding coefficients.

The square root of the first canonical correlation shows the proportion of the variance in the first canonical variate of one set of variables explained by the first canonical variate of the other set of variables. For the first canonical varite the proportion of explained variance is 61%.

The canonical redundancy analysis shows that neither of the first pair of canonical variables is a good overall predictor of the opposite set of variables, the proportions of variance explained being 0.24 and 0.35 for evaluation of the course and evaluation of the teacher respectively.

A four-dimensional structure of association between student evaluation of the course and student evaluation of the instructor can be a signal of data over-fitting due to an insufficient sample size. Another problem is that correlations between the variables within Form A and Form B are quite high.

Therefore, the CCA results are hard to interpret. Dimension reduction methods such as regularized and sparse versions of canonical correlation analysis should be used to obtain easier interpretable results.

5.3 Regularized CCA Results

The regularization was achieved by adding to the correlation matrices a corresponding identity matrix multiplied by a regularization parameter as described in the methods section. The cross-validation procedure was used to find the optimal regularization parameters. Only the first canonical correlation, equal to 0.70, appeared to be statistically significant ($p - value = 0.025$). Thus, this canonical correlation structure has only one dimension, compared to the four-dimensional result of classical CCA. This results in a simpler and more generalizable model of the association between evaluation of the course and evaluation of the teacher.

The interpretation of the results of regularized canonical correlation analysis is similar to the interpretation of the results of classical CCA. To investigate the structure of the canonical correlation, the standardized canonical coefficients and the structure canonical coefficients (canonical factor loadings and canonical factor cross-ladings) reported in Table 6 should be analyzed.

Table 6: Canonical weights and structure coefficients.

Evaluation of the course				Evaluation of the teacher			
	(1)	(2)	(3)		(1)	(2)	(3)
A.1.1	-0.08	-0.81	-0.57	B.1.1	-0.53	-0.96	-0.75
A.1.2	-0.03	-0.72	-0.51	B.1.2	-0.33	-0.93	-0.69
A.1.3	-0.33	-0.82	-0.62	B.1.3	-0.10	-0.88	-0.63
A.1.4	-0.03	-0.49	-0.34	B.2.1	-0,11	-0.58	-0.42
A.1.5	-0.38	-0.83	-0.65	B.2.2	-0.03	-0.82	-0.58
A.1.6	0.05	0.28	0.21	B.2.3	-0.03	-0.66	-0.46
A.1.7	0.03	-0.12	0.05				
A.1.8	-0.29	-0.85	0.65				

(1) - Standardized canonical weights; (2) - Canonical factor loadings; (3) - Canonical cross - loadings.

The analysis of the standardized canonical weights shows that questions A.1.3 (teaching material is good), in addition to A.1.5 and A.1.8 seen in the classical CCA, from the course related questions are the most important variables. Among the teacher related questions, B.1.1 (teacher gives me a good grasp of the academic content) and B.1.2 (teacher is good at communicating the subject) are the most important. An analysis of the canonical factor loadings and the cross-loadings shows that A.1.1 and A.1.2 from Form A and questions B.1.3 and B.2.2 from Form B also contribute to the canonical correlation.

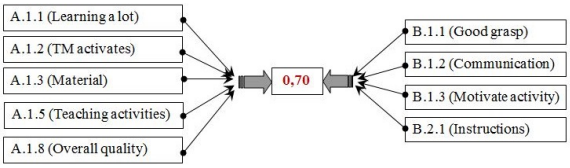


Figure 5: RCCA: Questions that contribute to canonical correlation.

Figure 5 presents the variables from Form A and Form B that contribute to the latent canonical variables.

An overall conclusion that can be made is that the correlation of 0.70 in the "Introductory Programming with Matlab" course is mainly due to the relationship between the content of the course, the teaching methods, the continuity between teaching activities in the course, the teaching material and the overall quality of the course from one side and the teachers ability to give a good grasp of the academic content of the course, the teachers ability to motivate the students, the teachers communication about the subject and the understanding of practical assignments on the other side.

5.4 Sparse CCA Results

The first canonical correlation of the sparse CCA was found to be equal to 0.75, which is the correlation between a linear combination of 4 variables from Form A and a linear combination of 3 variables from Form B. Table 7 presents the coefficients that correspond to these linear combinations.

Table 7: Sparse canonical coefficients.

Evaluation of the course		Evaluation of the teacher	
question	coef.	question	coef.
A.1.1	-0.08	B.1.1	-0.94
A.1.2	0	B.1.2	-0.32
A.1.3	-0.17	B.1.3	-0.14
A.1.4	0	B.2.1	0
A.1.5	-0.83	B.2.2	0
A.1.6	0	B.2.3	0
A.1.7	0		
A.1.8	-0.53		

From Form A, the questions: A.1.1, A.1.3, A.1.5 and A.1.8 contribute to the course related latent canonical variable while from Form B, the questions: B.1.1, B.1.2 and B.1.3 contribute to the teacher related latent canonical variable. This model is also simpler than the one obtained from classical CCA. Furthermore, it also involves less variables than the model obtained from the regularized version of CCA (it does not contain questions A.1.2, A.1.4, A.1.6, A.1.7, B.2.1, B.2.2, and B.2.3).

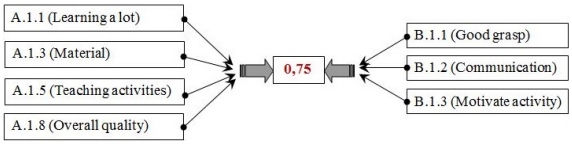


Figure 6: SCCA: Questions that contribute to canonical correlation.

Figure 6 presents the variables from Form A and Form B that contribute to the latent canonical variables. The conclusion is that the canonical correlation of 0.75 in the "Introductory Programming with Matlab" course is mainly due to the relationship between the good continuity between teaching activities in the course, content of the course, teaching material and overall quality of the course from one side and teachers ability to give a good grasp of the academic content of the course, teachers ability to motivate the students and teachers good communication about the subject on the other side.

5.5 Stability of the Results

To check for the stability of the correlation structures, subsequent years of the course should be analyzed. Figure 7 and figure 8 present the canonical correlation structures for the association between evaluation of the course and evaluation of the teacher in June 2011 and June 2012, respectively.

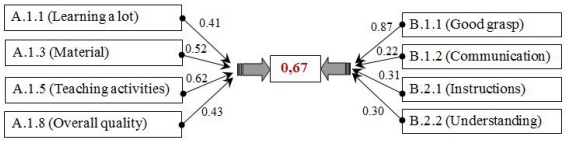


Figure 7: The structure of canonical correlation between the two parts of course evaluation in June 2011.

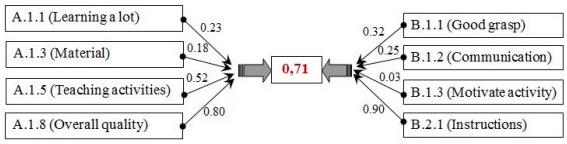


Figure 8: The structure of canonical correlation between the two parts of course evaluation in June 2012.

Overall, the two structures are similar. The only difference was on the evaluation of the teacher, where question B.2.2 (teacher's help to understand the academic content) form the structure in 2011, while B.1.3 (The teacher motivates us to actively follow the class) was in the canonical correlation structure for 2012. Figures also show the weights, each variable had in the latent canonical variable. The weights were different in the two years. However the changes in the canonical correlation structures can be explained by the fact that the main teachers of the course for all three semesters (January 2010, June 2011 and June 2012) were different.

6 DISCUSSION

The study have found that association between how students evaluate the course and how they evaluate the teacher of the course is strong (correlation is around 70 %), and the structure of this association is relatively stable over time. The square root of the first canonical correlation shows the proportion of the variance in the first canonical variate of one set of variables explained by the first canonical variate of the other set of variables (around 50%).

Having this strong relationship, better courses and therefore better SET results can be achieved in several different ways: improvement in a course can lead to better evaluation of teacher, and improvement of the teacher qualities, can lead to better evaluation of the course. However, Marsh (2007) indicated that students primarily evaluate the teacher rather than the course. But there is still some 30% left of this dimension, and there are the orthogonal dimensions as well.

The canonical redundancy analysis for the traditional CCA shows that neither of the first pair of canonical variables is a good overall predictor of the opposite set of variables, the proportions of variance explained being 0.24 and 0.35 for evaluation of the course and evaluation of the teacher respectively. There is no guarantee that what students answer on the course evaluation is not a function of the teacher, but the students do have some parts of non-correlating responses for the two parts of evaluation although there is also a strong association.

In case the structure of association between evaluation of the course and evaluation of the teacher is stable, SET administrators might consider to reduce the number of questions in the questionnaire in order to gain better response rates. However, that should be done very carefully. SETs must be multidimensional, in order to reflect multidimensionality of such a complex activity as teaching. According to Marsh and Roche (1997) , the strongest support for the multidimensionality of SETs is based on the nine factors: Learning/Value, Instructor Enthusiasm, Organization/Clarity, Group Interaction, Individual Report, Breadth of Coverage, Examinations/Grading, Assignments/Readings, and Workload/Difficulty. The questionnaire, currently used at DTU is already small, but an analysis similar to this could be used by other educational institutions.

7 CONCLUSIONS

This study analyzed the association between how students evaluate a course and how students evaluate a teacher using canonical correlation analysis (CCA). Data from student evaluations is characterized by high correlations between the variables within each set of variables, therefore two modifications of the CCA method; regularized CCA and sparse CCA, together with classical CCA were applied to find the most interpretable model of association between the two evaluations.

The association between how students evaluate the course and how students evaluate the teacher was found to be quite strong in all three cases. However, applications of regularized and sparse CCA to the present student evaluation data give results with increased interpretability over traditional CCA.

The simplest model was obtained from sparse canonical correlation analysis, where an association between how students evaluate the course and how students evaluate the teacher was found to be due to the relationship between the good continuity between teaching activities in the course, the content of the course, the teaching material, and the overall quality of the course from the course side; and teachers ability to give a good grasp of the academic content of the course, the teachers ability to motivate the students and the teachers good communication about the subject on the teacher side.

Analysis of subsequent evaluations of the same course showed that the association between how students rate the teacher and the course was found to be subject to subtle changes with the change of teaching methods and with the change of instructor. These changes in the correlation structure were seen on the instructor side and not on the course side.

REFERENCES

Abrami, P. C., d'Apollonia, S., and Rosenfield, S. (1997). The dimensionality of student ratings of instruction: what we know and what we do not. *Perry, R.P., Smart J.C., editors: effective teaching in higher education: research and practice. New York: Agathon Press*.

Cohen, P. A. (1981). Student rating of instruction and student achievement. *Review of Educational Research*, 51(3):281–309.

Déjean, S. and González, I. (2009). Package "CCA: Canonical correlation analysis". *CRAN*.

Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher education*, 30(6).

González, I., Déjean, S., Martin, P. G. P., and Baccini, A. (2009). Highlighting relationships between heterogeneous biological data through graphical displays based on regularized canonical correlation analysis. *Journal of Biological Systems*, 17(2):173–199.

Hardoon, D. R. and Shawe-Taylor, J. (2011). Sparse canonical correlation analysis. *Machine Learning*, 83(3):331–353.

Hoerl, A. and R.W., K. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67.

Hotelling, H. (1935). The most predictable criterion. *Journal of Educational Psychology*, 26:139–142.

Le Cao, K.-A., Martin, P. G. P., Christele, R., and Besse, P. (2009). Sparse canonical methods for biological data integration: application to a cross-platform study. *BMC Bioinformatics*, 10:Article 34.

Leurgans, S., Moyeed, R., and Silverman, B. (1993). Canonical correlation analysis when the data are curves. *Journal of the Royal Statistical Society B*, 55(3):725–740.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140:155.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3):253 – 388.

Marsh, H. W. (2007). Students evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. *R.P. Perry and J.C. Smart (eds.), The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*, pages 319 –383.

Marsh, H. W. and Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52(11):1187.

Parkhomenko, E., Tritchler, D., and Beyene, J. (2007). Genome-wide sparse canonical correlation of gene expression with genotypes. *BMC Proceedings*, 1.

Vinod, H. D. (1976). Canonical ridge and econometrics of joint production. *Journal of Econometrics*, 4(2):147–166.

Waaijenborg, S. and Zwinderman, A. (2007). Penalized canonical correlation analysis to quantify the association between gene expression and dna markers. *BMC Proceedings*, 1(Suppl 1):S122.

Witten, D. M., Tibshirani, R., and Gross, S. (2013). Package "PMA": Penalized multivariate analysis.

Witten, D. M., Tibshirani, R., and Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostat*, 10(3):515–534.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301320.

Reflections on the Use of a Personal Response System (PRS) for Summative Assessment in an Undergraduate Taught Module

Alan Hilliard
School of Health and Social Work, University of Hertfordshire, College Lane, Hatfield, U.K.
a.p.hilliard@herts.ac.uk

Keywords: Computer-aided Assessment, Assessment Software Tools, Technology Enhanced Learning.

Abstract: This paper outlines the author's experience of using a Personal Response System (PRS) for summative assessment in a 3rd year undergraduate taught module, over a 2 year period. The rationale for implementation of this method of assessment was a relatively high failure rate in the previous written examination (37%), and to reduce the marking burden for the teaching team. Key challenges identified with the implementation of the assessment process were reliability of the hardware/software, and student and staff confidence with the PRS and assessment process. Following the introduction of the new assessment method, the assessment failure rate was reduced to 9%. The PRS was seen as a good tool for summative assessment and received very positive student feedback comments. The PRS proved to be reliable, and with support and guidance, both students and staff felt confident with the process.

1 INTRODUCTION

Comparative Imaging is a 3rd year undergraduate taught module (30 credits at level 6), which explores the applications of different imaging modalities for three anatomical areas: head and neck, thorax and heart, and abdomen and pelvis. Previously the assessment of the module was by an individual written assignment, and a two hour unseen examination at the end of the module. In 2011, the examination was replaced by three, one hour long, summative multiple choice question (MCQ) tests, which were undertaken using the TurningPoint™ personal response system (PRS). The tests were staged throughout the module, each of the anatomical regions being assessed individually on completion of the taught content. Part of the rationale for the change, particularly with respect to staging of the assessment, was as a result of previous student feedback which highlighted that students felt overwhelmed with the volume of information learned, and a relatively high failure rate (37%) in the previous year's examination. The change in assessment method was also introduced to reduce the marking burden for the teaching team. This was significant, as there is a relatively large student cohort (125 students were registered on the module). The student cohort was also diverse with respect to age, with approximately 50% of the cohort being mature entry students, and also diverse with regard to ethnic and cultural backgrounds.

Previous research by the author investigating the use of the PRS for formative feedback with students had been very positive. 98.5% of students reported that the system was easy to use and 92.5% perceived that the use of the system helped their learning (Lorimer and Hilliard, 2007). Other studies have reported similar findings (Jefferson and Spiegel, 2009; Chen and Lan, 2013). Despite many studies exploring the use of PRS for formative feedback, there is little information on the use on the use of PRS for summative assessment.

2 USING THE PRS FOR SUMMATIVE ASSESSMENT

PRS typically comprise four elements: a tool for presenting lecture content and questions (e.g. a computer, PowerPoint™, and a digital projector), electronic handsets that enable students to respond to questions, a receiver that captures students' responses and PRS software collates and presents students' responses (Kennedy and Cutts, 2005).

In order to use the PRS for summative assessment, a number of steps needed to be undertaken: Firstly, a "participant" list needs to be created which includes the students' individual identification numbers along with a designated keypad number for each student. The student cohort

undertaking the assessment did not have possession of their own individual handsets, so an assessment signing in sheet was created. When students signed in at the assessment room, they showed their student identification to the invigilator who could then locate and give them their individual handsets. At the end of the assessment, the handsets were collected from the students prior to them leaving the assessment room.

Each test consisted of 20 MCQs, and each question was allocated a time limit of 2 minutes. Traditionally, when used for instant formative feedback, it is common to display a graph showing the percentage responses for each option answer, and to display the correct answer for each question. As the system was going to be used for summative assessment, the author felt that displaying this data would be a distraction for the students, and could impact negatively on their confidence if they saw that they were getting a number of questions wrong. Therefore, after each question was polled for 2 minutes, signalled by a countdown indicator, no data was displayed in the assessment room and the lecturer would move on to the next question.

Following the completion of each MCQ test, The PRS software was used to generate reports of individual student marks, and also generic feedback for each question with pie charts showing the spread of responses and identifying the correct answers. For each of the tests, the students' individual marks and generic feedback was made available through the module website on the University's managed learning environment (MLE) within a 24 hour period.

3 CHALLENGES

One if the primary considerations in considering adopting the PRS as an assessment method in the module, was the overall contribution of the assessment towards the students final module mark. The three PRS assessments were equally weighted at at 20% so, collectively, The PRS assessments contributed 60% towards the students' final module mark. The other 40% contribution towards the students' final module mark was by submission of an individually written assignment. As the PRS assessments contributed the greater portion of the Students final module mark, it was important to ensure that the assessment method and process were robust and reliable. Prior to the implementation of the PRS for summative assessment, a number of key challenges were identified:

i. Reliability of hardware/software.
ii. Student trust in the PRS and assessment process.
iii. Lecturer confidence in the PRS and the assessment process.

Similar challenges have been identified by other researchers (Roe and Robinson, 2010).

The PRS has been used extensively since 2005 for formative feedback, and during that time, had shown itself to be reliable. In order to be assured of reliability of its use for summative assessment a number of checks were put in place. Once the MCQ test had been created, it was given a trial run using a small number of handsets (approximately 6). This allowed a visual check of the questions for visibility and clarity of format. It also allowed for visibility and consistency of the countdown indicator on each question slide to be checked as a means of showing the open polling window and automatically closing the question poll after the allotted timeframe had elapsed. Finally, a report was generated to demonstrate that the question data had been captured and could be displayed as both individual marks and generic feedback.

A key challenge to the assessment method would be the need to gain the students trust and confidence in the process, particularly as this would be the first time that they had encountered this method of summative assessment. This was approached in a number of ways. Firstly, the assessment process was described in detail during the module induction session, which helped to demonstrate to the students that the structure and process of the assessment had been fully considered prior to its introduction. Secondly, early in the module following the first topic, there was a formative session arranged with a "mock" of an actual summative assessment. At this point handsets had been assigned to individual names, and were distributed to the students. The first slide asked students to "press button 1" to check that the handsets were working. During the polling timeframe, the "show response grid" button was selected on the TurningPoint™ toolbar. This projected a grid of all the handsets, and students could see their name change colour as they pressed their handset buttons (see figure 1 below – the student names have been removed to ensure confidentiality). This helped to increase the students' confidence and trust in the reliability of the technology.

Following the "mock" PRS assessment, the process was discussed and any questions or concerns were addressed. One issue that the students raised was concerns that their handset wouldn't work on

the day. It was agreed that a paper "receipt slip" would be used during the assessment and handed in with the handsets at the end of the assessments. The paper "receipt slip" listed the question numbers and students could write their answers in the boxes provided.

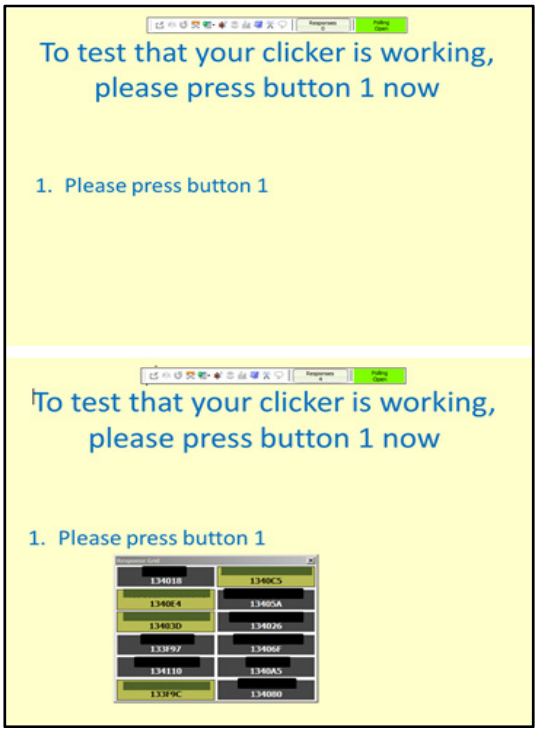


Figure 1: Showing the student participant response grid.

In adopting this system, three points were made clear. Firstly, the paper "receipt slips" would not be marked. Secondly, the paper "receipt slips" would only be considered where there was clear evidence of a non-functioning handset, and thirdly, because students could change their answers to questions within the polling window timeframe, there would be no discussion where students selected one answer using their PRS handset and wrote a different answer on their "receipt slip". The electronic response would be taken as the official answer.

Finally, the issue of lecturer confidence in the PRS and assessment process was not without its challenges. I had a number of years of experience of using the PRS for formative assessment, but had not previously used it for summative assessment. I was very conscious of the need for the process to work without problems to establish the students' trust from the outset. I would be the main administrator for the assessment, and a colleague would simultaneously undertake the assessment in a separate room with students who had specific study needs agreements. An experienced teaching team member was selected, and a step-by-step guide for the assessment process was produced and it was discussed prior to each assessment. Also before each assessment, every handset was checked that it was working and was set to the correct radiofrequency (RF) channel for the RF receiver so that data responses would be received. This was checked by again using the "show response grid". Next the assessment question slides were run using a small number of PRS handsets, and a report was generated to ensure that the software was working correctly.

4 STUDENT RESULTS AND FEEDBACK

Combining the marks for the individual PRS assessments resulted in an overall fail rate of 9%, as compared with the previous failure rate of 37%. Student feedback on the module was very positive:

"This was a very well organised module, and the modular assessment was helpful. This has been an enjoyable module to study".

"I really liked the structure of the in class tests... Got feedback of exam results the next day! All the teachers are really good and helpful".

"The current set-up of the module is great. I wouldn't change anything".

"This has been an excellent module…I like the way the module has been split into 3 sections which are tested throughout the semester. The PRS system of testing appears to work well and gives the student prompt feedback".

One feedback point that was mentioned by a number of students was about the linearity of the assessment, i.e. each question slide was answered in turn, without the opportunity for students to go back and review their answers to previous questions, which they would have liked. This point was taken on board for the second year that the assessment method was used. In order to enable this, students had a conventional paper-based MCQ. At the end of the allotted time, the PRS was used to collect the data. The answers to each of the questions were collected in turn, with students entering the answer they had previously written down. At the end of the assessment, the students handsets were collected in along with their paper-based MCQs. Here the students MCQ papers replaced the previous idea of the "receipt slip", and it was again made clear that the students' electronic responses would be taken as

their official answers.

5 PERSONAL REFLECTION ON THE ASSESSMENT PROCESS

Following undertaking this method of assessment over a 2 year period, and reflecting on the process, a number of advantages and challenges have been identified.

Some of the clear advantages of this method are the reduction in staff marking time, along with the ability to provide rapid student feedback and rapid release of marks. In each of the student assessments, individual marks and general feedback was provided within a 24 hour timeframe. An unexpected finding was that the assessment method and process appeared to be very suited to the ethnically, and culturally, diverse student cohort. Some students who had not performed well with traditional written examinations appeared to perform better with the PRS assessments. This finding is not specifically supported within current literature.

The process was not without its challenges, however. There was a considerable time input required to construct MCQs at the appropriate level, and which would challenge the students to demonstrate their level of knowledge and understanding. Additionally, time was required to create participant lists, check handsets and TurningPoint™ question slides prior to each assessment.

Due to the cohort size and the centralised timetable booking system at my institution, it became necessary to run the summative assessment in more than one room at the same time. This led to the need to train other teaching team members in the use of the PRS and how to conduct the summative assessment process. There was also a greater staff input required to invigilate the assessment in multiple rooms.

Any concerns regarding the reliability of the hardware and software, proved to be unfounded. Each of the assessment sessions ran smoothly and without technical error.

Student trust in the system and assessment process was established early in the module, through careful explanation, dialogue and discussion. It was supported by running a formative "mock" assessment, which enabled students to experience the process and become familiar with using the PRS handsets. Student feedback on the module provided further evidence regarding the success of this method of assessment.

Lecturer confidence in the system and process could be a challenge if staff felt inexperienced in the use of the PRS or lacked understanding of the structure and process of the assessment. Myself, as lead lecturer, had considerable experience in the use of the PRS and took responsibility for the design and structure of the assessment. Where other teaching team members were required to administer the assessment, support, training and guidance were given. The teaching team managed to successfully run the assessments, and reported high levels of confidence as a result.

6 CONCLUSIONS

The PRS has shown itself to be a successful tool for summative assessment, which can save on staff marking time, although there is a need to be mindful of the time requirement to set up and check the PRS and handsets prior to use. The assessment method may be well suited for ethnically and culturally diverse suited cohorts, although this has not been established. It is important to support students, so that they have trust and confidence in the assessment process and the use of the PRS. Similarly, it is important to recognise the needs of the staff teaching team, and to provide training, support and guidance as required.

REFERENCES

Chen, T., Lan, Y. 2013. Using a personal response system as an in-class assessment tool in the teaching of basic college chemistry. *Australasian Journal of Educational Technology*, 29(1), 32-40.

Jefferson, W., Spiegel, D. 2009. Implementation of a university standard for personal response systems. *Association for the Advancement of Computing in Education Journal*, 17(1), 1-9.

Kennedy, G. E., Cutts, Q. I. 2005. The association between students' use of an electronic voting system and their learning outcomes. *Journal of Computer Assisted Learning*, Vol. 21, 260-268.

Lorimer, J., Hilliard, A., 2007. Net gen or not gen? Student and Staff Evaluations of the use of Podcasts/Audio Files and an Electronic Voting System (EVS) in a Blended Learning Module. Conference Proceedings: *6th European Conference on e-Learning*, 407-414.

Roe, H. M., Robinson, D. P., 2010. The use of Personal Response Systems (PRS) in multiple-choice assessment: benefits and pitfalls over traditional, paper-based approaches. *Planet*, Issue 23, 54-62.

Two New Concepts in Video Podcasts
Minimalist Slides and Modular Teaching Mini-videos

Emilio Letón[1] and Elisa M. Molanes-López[2]
[1]*Department of Artificial Intelligence, UNED, Juan del Rosal 16, 28040, Madrid, Spain*
[2]*Department of Statistics, Carlos III University of Madrid, Avda. Universidad 30, 28911, Leganés (Madrid), Spain*
emilio.leton@dia.uned.es, elisamaria.molanes@uc3m.es

Keywords: Learning Resources, Minimalist Slides, Mini-videos, MTM, Rubrics, Video Podcasts.

Abstract: Nowadays the main learning resources are traditional board, slides and video podcasts. Each of them with their own strengths and weaknesses. Analysing the advantages and disadvantages of the traditional board and slides we have created, with the advantages of both resources, a new concept that we have named "minimalist slides". Using the minimalist slides we have introduced a new type of video podcast that we have named "modular teaching mini-video".

1 INTRODUCTION

In today's teaching there are different methods of transmission of knowledge by the teacher: traditional board, slides and video podcasts. Traditional board and slides are well known and are still used. Video podcasts as a learning resource are being used more and more nowadays (Kay, 2012, among others). In many cases, video podcasts are just the recording of entire lectures that are viewed passively (Traphagan et al., 2010; Wieling and Hofman, 2010). Other forms of video podcasts are shorter in time and are associated with interactive tasks to make the learning more effective. For instance, Griffin et al. (2009) used synchronized PowerPoint and voice and He et al. (2012) and Loomes et al. (2002) voice and handwriting.

Each of these methods has strengths and weaknesses. In this work, we analyse the advantages and disadvantages of the traditional board and slides with a survey of students at a Madrid (Spain) university to create, with the advantages of both, a new concept that we have named "minimalist slides" (MSs). With the help of MSs we have created a new type of video podcast that we have called "modular teaching mini-video" (MTM).

The outline of the paper is as follows. In section 2 we analyse the advantages and disadvantages of the traditional board and slides. In section 3 we introduce the MSs and in section 4 the MTMs. In section 5 we propose several rubrics. Finally, we describe some examples of MTMs in section 6.

2 TRADITIONAL BOARD AND SLIDES

In this section, we analyse the strengths and weaknesses of the traditional board and slides with a pilot survey of 18 students studying Statistics in the second year of Engineering at Madrid's Carlos III University. The survey had open questions. We show below similar answers given by the students to each question to have a better general view. Additionally, they were asked to allocate 100 percentage points in total among three methods (blackboard, slides, mixture) according to their preferences for a theoretical class. The results were as follows: a median of 44% for blackboard, 38% for slides and 18% for a mixture.

2.1 Traditional Board

The advantages of the traditional board from the students' viewpoint are:
- It is more dynamic, it forces you to follow the subject and you can customize your class notes.
- You can highlight or emphasize text. The explanations become clearer.
- You can copy the concepts as the teacher explains and thus you have time to understand and make your own notes.
- We have more time to copy. You can solve the problems as you would in an exam.
- It is easier to take class notes.

- Since you have to write, whether you like it or not, you will learn even if you pay little attention.
- The teacher explains more on the board and more of what is on the slides.
- As the teacher explains it is easier to see how the explanation evolves. In my opinion, you take better notes from the board.
- Since the teacher is writing, the student also has time to write in his/her notebook.
- It is less boring and everything is explained a little bit more slowly and you have more time to absorb the explanations.
- You have more possibilities of having all the notes because the teacher can go more slowly.
- There is interactivity. It is easy to give examples.
- It forces the teacher to give an explanation at a moderate speed, allowing the student to copy notes and understand at the same time.

The disadvantages of the traditional board from the students' viewpoint are:
- I would have liked more clarity on the board and a slower explanation.
- It may happen that you cannot understand the teacher's handwriting or that he/she writes too fast and you do not have enough time to copy.
- It slows down the class a bit.
- More copying is required and the pace is slower.
- The teacher has to clean the blackboard often and you cannot go back. The blackboard is messy.
- You may not understand the teacher's explanations because you are trying to copy them.
- The teacher can make mistakes on the blackboard, with illegible handwriting or especially 3D drawings which are difficult to see.
- The teacher is much more messy and disorganized. Sometimes it is difficult to understand what is written and you have to ask him/her.
- You are busier copying than understanding.
- More time is needed to present things such as tables with many data.
- It depends on the teacher's skills to explain himself/herself clearly, in an organized manner on the blackboard.
- It is more difficult to retain all the information if the subject is more elaborate.
- The main disadvantage, that is not always present, is the lack of understanding due to poor lighting conditions or unclear handwriting.

In summary, the strengths of the traditional board are that it brings dynamism, it is easier to take notes, the teacher explains with more examples, it is easier to follow the teacher's explanations because the pace is slower and there is more interactivity.

The weaknesses are that it slows the class a little, it is more difficult for the teacher to be organized, often he/she does not get a "clean" finish, it is difficult to copy and understand at the same time and the quality of the drawings is not so good.

2.2 Slides

The advantages of the slides from the students' viewpoint are:
- The transfer of knowledge is quicker because the learning content written on the slides is more clear and organized.
- Since you have the slides in advance, you can follow perfectly the explanations.
- The graphical plots are clearer. Besides, the slides provide support to the teacher's explanations and he/she can advance at a convenient speed.
- You do not have to copy.
- They are more visually attractive and facilitate reading. They are better suited for theory.
- They provide clear explanations that help us to learn better.
- Since you already have the slides, it is easier to add your own comments and to follow the teacher's explanations.
- The big advantage is that the quality of the 3D drawings is improved.
- They are easy-to-access material on the web.
- They are useful for us to study for the exams.
- They summarize the concepts and although sometimes there may be information missing that makes the understanding more difficult, though since they contain the most important concepts, you finish the class with a better understanding.
- They are a helpful support for the students when they miss a class.

The disadvantages of the slides from the students' viewpoint are:
- You are more lost because having the slides you write less.
- I do not like the classes with slides because the explanations are not given thoroughly enough.
- The slides are not as your own notes. Besides, you may lose concentration and not take notes of the things that are not included on the slides.
- Sometimes the slides are too summarized. There is missing content.
- Sometimes the content of the slides is not enough and the classes become tedious and too passive.
- If things are not cleared it is more difficult to understand them. Sometimes the explanations go too fast. You do not learn so much.
- As you do not have to copy, sometimes you do not pay attention to them.

- It is a too general and theoretical material when you are asked to solve exercises.
- The explanation is understood less easily than when the teacher explains on the board.
- It may happen that you do not have enough time to add your own comments.
- There is a tendency to go faster (can be annoying).
- The worst way of teaching is with slides.
- They can make the students sleep. Sometimes they are too schematic. It is more difficult to take notes.
- The slides seem to be a comfortable way of teaching for the teacher but very bad for the students because it is very difficult for them to take notes and the teacher usually goes too fast.

In summary, the advantages of the slides are that they are more organized and allow the teacher to go faster, the student does not need to copy, their finish is clean, they can be supplemented with multimedia and graphic support, and they are easily accessible on the web. The disadvantages are that they sometimes allow the teacher to go too fast, they favour a passive attitude, they are often too summarized, the explanations are more concise and there is no space to take notes.

3 MINIMALIST SLIDES

We have seen in the previous section that one advantage of the slides is that the student does not need to write because the slides contain all the information for the student ("maximalist slides"). But this advantage seems to be also a disadvantage for the students because they do not feel the need to write and it causes a passive attitude in them. Even though the teacher can give additional explanations by using his/her voice and/or writing on the board, the student may not pay attention because he/she assumes that explanations not written on the slides are never important. For that reason we propose to use slides with little information on them ("minimalist slides", MSs). MSs are not self-explanatory and force the teacher and the student to work with them. The MSs are used as follows: the teacher provides the MSs on the web previously to the class, the student prints them and takes the MSs to the class, the teacher projects the MSs and writes on them using a digital board to fill them out while the student is copying these notes on his/her printed version of the MSs. Due to the fact that some information is already written on the MSs the student has more time to think as he/she is writing the missing information that is not included on the MSs. This way of proceeding avoids a line being copied/read by the student without thinking, which would be useless.

The basic features of the MSs are:
- The number of lines on each slide is less than or equal to 7.
- There is space to write on each slide.
- The slides are available to be printed.

After using the MSs during a whole academic year in 40 students of the same subject, they were asked to allocate 100 percentage points in total among three methods (MSs, blackboard, maximalist slides) according to their preferences for a theoretical class. The results were as follows: a median of 80% for MSs, 1% for blackboard, and 19% for maximalist slides.

The MSs have the advantages of the traditional board and slides and do not have the disadvantages of them. The only disadvantage is that they cannot be used in distance learning. This disadvantage will be solved in the new section with the MTMs.

4 MODULAR TEACHING MINI-VIDEOS

In this section, we introduce a new type of video podcast that we have named modular teaching mini-video (MTM). This concept is characterized by its duration, support, methodology, philosophy, format, and interconnection, as explained below.

The duration of an MTM is between 5 and 10 minutes. In order to achieve this duration, an effort of synthesis has to be made and it is necessary to have a structured vision of its content (this is not difficult, but it requires some practice). Nevertheless, we have to remark that "not everything can fit" within an MTM. The MTM is a complement to other traditional teaching materials. The fact of its short duration makes it necessary to prepare the recording and rehearse several times. However, at the same time we have to be practical and to not worry about perfection.

The support in which an MTM is designed are the MSs, described in the previous section.

The methodology that is used in an MTM is based on the digital board, where the MSs are projected and filled in simultaneously. This way of proceeding prevents the passive attitude of students. While watching the MTM the student has to work on each line or drawing that is hand-written by the teacher on the slide. So, first the student has to think what must be filled in on the slide, and then he/she has to write it while he/she is watching the MTM.

Besides, on one hand, the speed of this process is the one that the student needs and, on the other hand, the digital finishing is the one that the student wants. These elements are achieved using the digital board.

The MTMs are based on the following underpinning philosophy: I work (the teacher works) / you work (the student works). We can ask ourselves if the actual teaching will have success or not, both in distance learning and in the classroom (there are not many differences between them: in fact, they are converging). To answer this question we can think of four scenarios, whether the teacher works or not and whether the student works or not. There is only one of these scenarios, where both work, that will have success in the actual teaching. In this scenario, the teacher makes educational material and activities (for example, MTMs) and the student works with them. This philosophy has relevant advantages for the student and the teacher.

The format in which you can view an MTM is on the web ("streaming") or on a mobile telephone or similar devices (3GP, MP4, etc.). Nevertheless, it is worthwhile to point out that the optimal way of the first viewing is on the computer after printing the MSs, in order to fill them out, following the pace of the digital board. On a second viewing, mobile devices are ideal for reviewing, and this can be done in any place (for example, going to work or to university). With the duration of an MTM being so short, an easy computer processing and an easy search of contents is possible. Besides, there is a real possibility of interchange between students through "bluetooth". It is necessary that the teacher chooses an adequate size of fonts and elements in an MTM to take into account that the MTM has to be perfectly seen on mobile devices even with small screen sizes. It cannot be said: "Excuse me, I do not know if this line can be perfectly seen".

The interconnection (or modularity) refers to the fact that the design of an MTM has to take into account that the MTM could be used in several subjects (introductory courses at university, undergraduate or postgraduate levels) taking care also of the accessibility. The short duration of MTMs facilitates their modularity, in contrast to traditional lectures whose long duration do not help to reuse them.

5 RUBRICS

In this section, we propose several rubrics to evaluate the degree of minimalism of the slides that are used in any audio-visual material and the degree of MTM of any audio-visual material.

5.1 Minimalist Slides

The rubric to assess the degree of minimalism of the slides that are used in any audio-visual material is to give a mini-point for each of the items listed below:

1. The number of slides is less than or equal to 10.
2. The number of lines on each slide is less than or equal to 7.
3. There is space for subtitles.
4. There is space to write on each slide.
5. The upper right corner of each slide is reserved for the small image of the teacher.
6. The bottom right corner of each slide is reserved for the big image of someone doing sign language.
7. There is a slide to make a summary.
8. There is a front cover slide and a back cover slide.
9. The slides are available to be printed.

The final score is to add up these mini-points, multiply by (10/9) to get a 0-10 number. The slides are considered MSs if the final score is bigger than or equal to 8 and scoring at the first and last item.

5.2 Modular Teaching Mini-videos

The rubric to assess the degree of fulfilment of the MTM features of any audio-visual material is to give a mini-point for each of the following items:

1. It is between 5 and 10 minutes.
2. A synthesis effort is observed.
3. The slides are minimalist.
4. There is space for subtitles.
5. The teacher writes on the minimalist slides.
6. You can see well what the teacher writes.
7. A question is formulated that later on is solved.
8. A summary at the end is requested.
9. The written text can be clearly read on a mobile device.
10. The text that the teacher writes at that moment can be clearly read on a mobile device.
11. It can be used in several subjects.
12. The subtitles are available.

The final score is to add up these mini-points, multiply by (10/12) to get a 0-10 number. The audio-visual material is considered an MTM if the final score is bigger than or equal to 8 and scoring at the first and last item.

6 EXAMPLES

In this section, we give some examples of MTMs. To follow completely the explanations on them, it is

required to have a mathematical knowledge equivalent to a student of approximately 14 years old. Nevertheless, the MTMs can be applied to other non-mathematical fields as is indicated in the following subsections. The examples of the MTMs described in this section and their corresponding MSs are available at www.ia.uned.es/minivideos.

6.1 Characteristics of an MTM

In this MTM, the characteristics of an MTM are described in terms of its duration, support, methodology, philosophy, format, and interconnection. In Figure 1a, an "empty" MS (without being filled out) is shown with the characteristics of an MTM. The same MS but "filled in" is shown in Figure 1b at the moment of the recording of the MTM. This MTM shows how to introduce theoretical concepts using an MTM. In this case, having in mind that these concepts were the theoretical concepts of an MTM and they have been explained in MTM format, we could say that it is a meta-MTM. Other examples could be:
- Advice to write (tales, short notes...).
- Hints to enjoy art (painting, cathedrals...).
- Use of the social networks in Education.
- Types of literary works, musical periods...
- The importance of music in films.
- What are the origins of Philosophy?

6.2 False Proof I

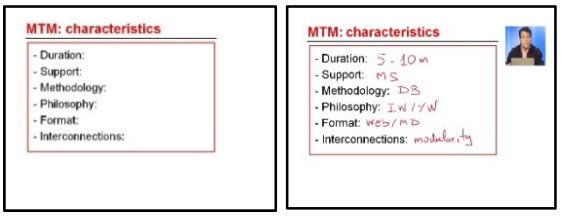


Figure 1: a) Empty MS: characteristics of MTMs (left). b) Filled MS: characteristics of MTMs (right).

In this MTM, one must pay attention to discover the moment when an erroneous reasoning is made in a "proof". At the end of the MTM the conclusion is an absurdity, specifically that 2=1. If we are not able to see where the reasoning is incorrect, we will have to accept this absurd conclusion and that would not be reasonable. This is an example of MTM where an erroneous mathematical reasoning is made (although apparently right) to check if the student is able to detect that error. This same idea can be used in other fields. For instance, detect errors when:
- Using languages (English, Latin...).
- Analysing a sentence.
- Classifying a master piece (artwork, music, ...).

6.3 Second Degree Equation

In this MTM, the formula to solve a second degree equation is deduced. This formula is easy to remember and easy to deduce. In the following second degree equation, $ax^2 + bx + c = 0$, it is necessary to assume that $a \neq 0$ as explained in the filled out MS of Figure 2b, associated to the empty MS of Figure 2a, because otherwise the equation would be of first degree rather than of second degree, given by $bx + c = 0$. In Figure 2b, it is indicated that one cannot divide by 0 as has been illustrated in the MTM entitled "False proof I". This is an example of MTM connected to another (in this case, the MTM entitled "False proof I"). This same idea could be applied in other fields, for instance:
- Latin/English in the Music (Enya, Carl Orff... / The Cranberries, The Smiths...).
- Advice to make text commentaries.

6.4 John and Paul

In this MTM, we are going to solve a very simple riddle about the number of sheep that the shepherds John and Paul have. The riddle is as follows:
I) John tells Paul: "If you give me a sheep, I'll have twice as many sheep as you".
II) "However, if I give one to you, we will have the same number of sheep".
How many sheep had John at the beginning?
The statement I) is schematically expressed in the empty MS of Figure 3a and its algebraic expression in the filled out MS of Figure 3b. The statement II) is schematically expressed in the empty MS of Figure 4a and its algebraic expression in the filled out MS of Figure 4b. This riddle is solved mathematically in the MTM. However, the purpose of this MTM is to show that it is possible to solve a mathematical exercise through several methods (formal vs. intuitive) and that it is important to choose a good notation to do so. This idea could be

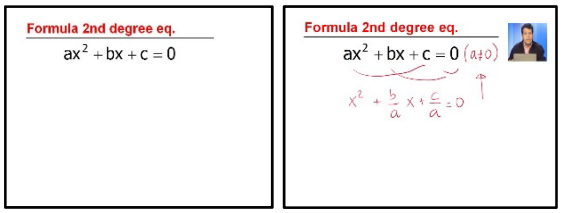


Figure 2: a) Empty MS with first steps of the MTM (left). b) Filled in MS with first steps of the MTM (right).

applied in other areas, for instance:
- To show several solutions of practical cases of criminal law, properly argued.
- To propose several algorithms to solve a specific problem, either mathematical, computational or from the everyday life.
- To implement the same pseudo code using different languages of programming.

6.5 Three Daughters

In this MTM, we solve a very simple riddle although perplexing at first sight. It is about two childhood friends and the ages of the three daughters that one of them has. The riddle is as follows: A man that has three daughters says to one of his childhood friends: "Let's see if you can guess the ages of my three daughters, with the following clues:
I) "The product of their ages is 36".
II) "The sum of their ages coincides with the number of the building where we used to wait for each other to go to school together"."
The friend starts thinking and after doing some calculations he says: "Something is missing." The other friend replies:
III) "Yes, you are right. The oldest is a pianist".

This riddle is solved in the MTM in an exhaustive way, listing all the possible cases for the ages of the three daughters. The objective of this MTM is to show that you must not surrender to a problem that seems intractable at first.

This idea can also be applied in other situations, for instance:
- To describe specific scientific enigmas and how different theories have tried to explain them.
- To solve crimes with the help of forensic science.

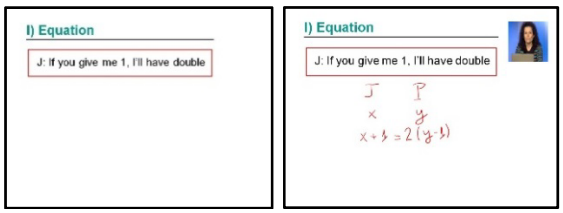


Figure 3: a) Empty MS with the statement I) (left). b) Filled in MS with the statement I) (right).

7 CONCLUSIONS

We have introduced the concept of "minimalist slides" (MSs) which incorporates the advantages of the traditional board and slides and which does not have their disadvantages. The MSs have been incorporated into a new type of video podcast

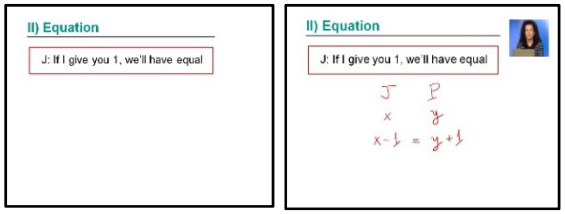


Figure 4: a) Empty MS with the statement II) (left). b) Filled in MS with the statement II) (right).

creating the "modular teaching mini-videos" (MTMs) that inherit the advantages of the MSs with the additional benefit that they can be used at any time and place that the student chooses. We believe that the MTMs will be very useful in the actual teaching covering classroom teaching, blended learning and e-learning. In order to prove this belief we are planning different experiments according to the several characteristics of the MTMs.

REFERENCES

Griffin, D. K., Mitchell, D., Thompson, S. J. (2009). Podcasting by synchronizing PowerPoint and voice: what are the pedagogical benefits? Computers & Education 53 (2), 532 - 539.

He, Y., Swenson, S., Lents, N. (2012). Online Video Tutorials Increase Learning of Difficult Concepts in an Undergraduate Analytical Chemistry Course. *Journal of Chemical Education 89 (9)*, 1128 - 1132.

Kay, R. H. (2012). Exploring the use of video podcasts in education: a comprehensive review of the literature. Computers in Human Behavior 28 (3), 820 - 831.

Loomes, M., Shafarenko, A., & Loomes, M. (2002). Teaching mathematical explanation through audiographic technology. Computers & Education, 38(1 - 3), 137 - 49.

Traphagan, T., Kucsera, J. V., Kishi, K. K. (2010). Impact of class lecture webcasting on attendance and learning. Educational Technology Research and Development 58 (1), 19 - 37.

Wieling, M. B., Hofman, W. H. A. (2010). The impact of online video lecture recordings and automated feedback on student performance. Computers & Education 54 (4), 992 - 998.

Academic Frontier-based Approach Based on Constructivism

Nuo Liu[1], Shengli Yang[2], Bing Feng[3] and Liangyin Chen[4]
[1]*School of Microelectronics and Solid State Electronics,*
University of Electronic Science and Technology of China, Chengdu, China
[2]*Zhejiang University of technology, Zhejiang Province, Hangzhou 310014, China*
[3]*School of Foreign Language, University of electronic science and technology of China, Chengdu, China*
[4]*College of Computer and Science, Sichuan University, Chengdu Sichuan, China*
liunuo2002@gmail.com, {yangshengli01@zjut, bfeng@uestc, chenliangyin@scu}.edu.cn

Keywords: Constructivism, Academic-Frontier Approach, Open-ended, Collaborative, Higher Education.

Abstract: For the purpose of combination of the basic and academic frontier knowledge by the application of the theory of constructivist learning in order to trace the rapid development of microelectronics, Academic frontier-based approach (AFA) is presented. The learner-centered instructional approach is valuable for promoting active learning by involving learners in learning academic frontier topics in an open-ended and collaborative environment. The design and implementation enrich the teaching modes and the content of the key curriculums in depth. It is effective in achieving positive and higher cognitive goals.

1 INTRODUCTION

Many Universities have a growing concern with the learners' failure rates and academic achievement in engineering courses. Learners and instructors are often skeptical of the effectiveness of this technique. Some of the frequent complaints that we hear from our learners is "how to apply the knowledge in the textbook", "the lecture-centered study is inactively and bored", and we "have not shown them how to do." According to Shuell (Shuell, 1986), "If learners are to learn desired outcomes in a reasonably effective manner, then the lecturer's fundamental task is to get learners to engage in learning activities that are likely to result in their achieving those outcomes". John Dewey, the father of progressive education supports 'learning by doing'. Tyler (Tyler, 1949) wrote, 'Learning takes place through the active behavior of the learner; it is what he does that he learns, not what the teacher does.' Constructivist theories have being the common theory-in-use in higher education since the centrality of the learner is given. For instructors, how to inspire those young and fresh minds found in academia? How to develop the learners' curiosity in the subject they will face after stepping off campus?

We present Academic Frontier-based Approach (AFA) in the paper. The work reflects our on-going research on teaching and learning in higher education in which we have developed in several of our engineering courses by learners' actively constructing new knowledge for the purpose of learners' academic pre-research and motivation.

2 WHAT IS ACADEMIC FRONTIER APPROACH?

Academic Frontier Approach **(AFA)** is a pedagogy that centers learner learning on Academic Frontier-driven projects facilitated by a instructor in order to achieve the new learning outcomes of courses related to the content in textbooks — expansive learning. Figure 1 gives the outline of AFA. It is a cognitive constructivist epistemology which concludes from science references that learners gain more through relating academic material and papers to their own interests and academic vision, and that such experience informs their ability to conceptualize content (Duffy and Jonassen, 1992). Constructivism calls for learning opportunities that are experiential, active, and collaborative. The goals here for the learner are: firstly, to develop and implement the new forms of teaching approach; secondly, cultivate the curiosity and motivation by expanding and collaborating learning in groups and

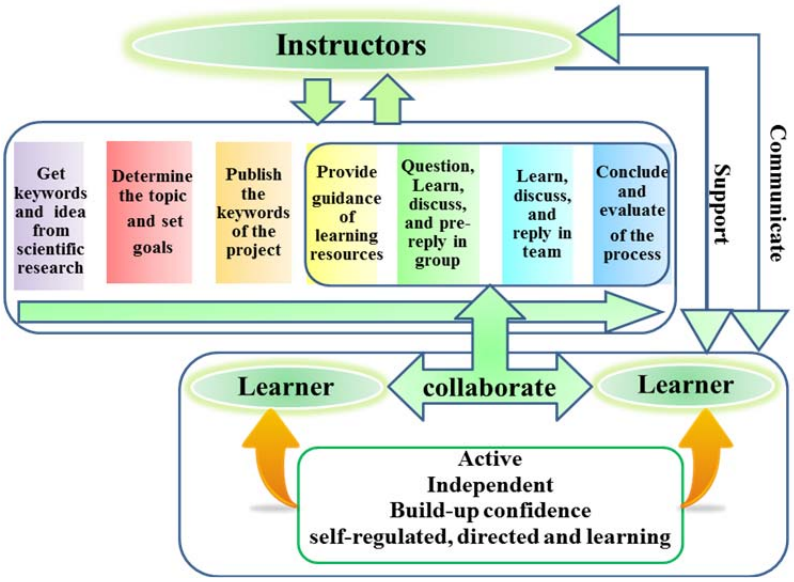


Figure 1: The outline of the academic frontier-based approach.

teams. Thus, it is not to passively absorb and regurgitate information; but rather to actively engage with the science references, through an inquisitional learning, discussing and analysis with group mates, and effectively review and get academic frontier knowledge with the corresponding basic knowledge gained in textbook. Therefore, the ultimate goal is the development of knowledge expanding, expansive learning, interest adding, and potential exploiting abilities which are beneficial for learner's careers in the future research and work.

For our work, we have found several underlying themes to be particularly cogent toward our instructional activities. Specifically, AFA provides a context for the content of our courses; meaning that many of the individual concepts that each course identifies are drawn together to understand the new knowledge that are submitted to learners through the various keywords and topics that they work on during their AFA project work. It is important then to draw the distinction that AFA presents specific topics for learners to learn. The advantages of AFA project works in microelectronic courses are the increased engagement of learners and the relationship to professional practice, team cooperation and self-regulated and directed. Thus it is common to assign keywords and topics as a beginning for learners to illustrate their abilities of self-learning and reviewing of science references by group work, and indeed to assess that learning under the guidance of instructors. The importance here is that we are not solely doing AFA project work, but also organize and construct advanced knowledge (expansive learning) based upon the well known content knowledge that has been written in textbooks. Such scaffolding is integral to the process as learners use the content provided along with their own learning with their groupmates/teammates, to construct new knowledge based upon the keywords and topics at academic frontier. Learners learnt in this manner would get an important experience in the process. Therefore, AFA and project-based learning (PBL) are similar in that there is a shared goal of successfully completing the activity in a learning community actively; however, it is different from the latter in that projects typically cultivate the abilities of review and analysis of references. Moreover, the implementation of AFA is a process of sharing information and goals, and improving social relationship by collaborative learning. Thus, AFA has the characteristics of PBL and collaborative/cooperative learning.

While AFA can be applied in any discipline, its appeal within microelectronic science is clear. The courses, such as "Semiconductor Physics", "Physics of Semiconductor Devices" or "Compound Semiconductor Materials and Devices", are courses designed to implement AFA. With the rapid advances within the microelectronic, it is also of particular concern that learners understand how to be

good independent learners that catch up with the progress of microelectronics. We will particularly value the pedagogy that results in graduates who are able to educate themselves about new technologies and conductor them AFA for their development in the future research and work. Thus, AFA directly supports the characteristics: an appreciation of the interplay between classic theories in textbooks and new knowledge at academic frontier, a familiarity with common themes and principles.

3 THE STRATEGIES AND PROCESS OF IMPLEMENTATION

3.1 The Constructive Strategy

(1) About knowledge

- To track the development of knowledge at scientific frontier related to textbooks.
- To increase new content of teaching information and the depth of the topic out of the textbook.

(2) About skills

- To co-integrate knowledge.
- To coordinate and grow up the teamwork of the learners.
- To cultivate the learners' independent, active, communicative, self-regulated and directed skills by collaborating learning.

3.2 The Process of Implementation

- Inspiration. Instructors ask questions; provide the keywords and topics about academic frontier.
- Introduction. Instructors introduce the academic background and the main website of the references.
- Definition of goals and formation of community of learning. Instructors allocate the task to a team which is consisted of 5-10 small groups according to the scale of the class. The teaching aim is decided and the significance of the topics is explained.
- Instruction. Instructors guide the learners to learn online resources and track the development of the academic.
- Modification of sub-goals. Learners learn, ask question, read, analysize, discuss and review by themselves and cooperate with the group-mates by QQ group and face-to face discussion in group meeting. Each group gives the presentation of the sub-goal.
- Integration of goals. The team gives the final presentation by integrating the key contents from each groups and discussion in class.
- Evaluation of the process by instructors.

4 CONCLUSIONS

To sum up, the AFA to teaching and learning has the characteristics as follows: The integration of knowledge of academic frontier and the essentiality of study; the exploration and the target guidance; the fun of teaching, the extensibility of teaching and effectiveness; construction; the cumulativity of the study; the diagnosis and the rethinking of learning; the society and collaboration.

ACKNOWLEDGEMENTS

The authors acknowledge the support from the International Corporation and Communication Scholarship of Sichuan Province (Grant No. 2012HH0027), Project of Undergraduate Education and Teaching Reform of UESTC, and Project of Undergraduate Education and Teaching Reform of ZJUT.

REFERENCES

Duffy, T. M, & Jonassen, D. H. (Eds.), 1992. *Constructivism and the technology of instruction: A conversation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Shuell, T. J., 1986. *Cognitive conceptions of learning, Review of educational research* 56, 411.

Tyler, R. W., 1949. *Basic principles of curriculum and instruction*, University of Chicago.

Using Technology to Accelerate the Construction of Concept Inventories

Latent Semantic Analysis and the Biology Concept Inventory

Kathy Garvin-Doxas[1], Michael Klymkowsky[2], Isidoros Doxas[3] and Walter Kintsch[4]
[1]*Center for Integrated Plasma Studies, University of Colorado, Boulder, CO, U.S.A.*
(Present address: Boulder Internet Technologies, Columbia, MD, U.S.A.)
[2]*Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO, U.S.A.*
[3]*Center for Integrated Plasma Studies, University of Colorado, Boulder, CO, U.S.A.*
(Present address: BAE Systems, Columbia, MD, U.S.A.)
[4]*Institute of Cognitive Science, University of Colorado, Boulder, CO, U.S.A.*
{garvindo, klymkowsky, doxas, walter.kintsch}@colorado.edu

Keywords: Concept Inventory, Biology Concept Inventory, Misconceptions, Latent Semantic Analysis.

Abstract: Concept Inventories are multiple choice instruments that map students' conceptual understanding in a given subject area. They underpin some of the most effective teaching methods in science education, but they are labour intensive and expensive to construct, which limits their wide use in instruction. We describe how we use Latent Semantic Analysis to accelerate the construction of Concept Inventories in general, and the Biology Concept Inventory in particular.

1 INTRODUCTION

Concept Inventories are multiple choice instruments that explore students' conceptual understanding in a given subject area. To accomplish this, CI developers look for verbal markers that can be used as proxies for identifying students' conceptual structures, much as we try to find DNA markers for various traits. Well constructed CIs provide researchers with a map of the students' conceptual landscape, which can be used to inform instruction in that area.

Research-based teaching methods that are firmly based on misconception research and make consistent use of collaborative learning are the most widely used national-scale tested methods that consistently produce learning gains significantly superior to lectures in Physics and Astronomy (eg. McDermott et al., 1998; Zeilik et al., 1997; Hake 1998). Short of one-on-one tutoring (cf. Bloom's "two sigma challenge", Bloom, 1984), this is the best model available for impacting student learning.

Although consistently successful, the model also incorporates a significant barrier to its wide adoption, replicability, and extensibility. It is critically dependent on the existence of well-researched assessment instruments that can reliably diagnose a student's misconceptions, and which require considerable time and effort to produce. Although several groups, both academic and commercial are currently engaged in developing such instruments in disciplines such as biology (e.g. Garvin-Doxas and Klymkowsky, 2008; Smith et. al., 2008; Kalas et. al. 2013), geoscience (e.g. Libarkin and Anderson, 2006), and engineering (e.g. Midkiff et. al., 2001), no substantial advance has been made in the time, effort, and expense required to develop a validated, reliable instrument.

Here we describe the construction of Concept Inventories and how it differs from the construction of tests, and we show how we use Latent Semantic Analysis (LSA, Landauer et. al., 1998; Landauer and Dumais, 1997) to facilitate the usually labour intensive validation phase of Concept Inventories in general, and the Biology Concept Inventory (Garvin-Doxas and Klymkowsky, 2008; Klymkowsky and Garvin-Doxas, 2008) in particular.

2 CONCEPT INVENTORIES AND TESTS

Although CIs bear a strong resemblance to standardized tests, the two types of instruments are very different, having fundamentally different aims. Tests are basically designed to answer the question "what percentage of the desired knowledge and skills in this field has this student acquired?". CIs are meant to answer the question "what conceptual constructs is this student using when solving problems in this field?". These same questions can also be asked from the point of view of an ensemble of students (rather than the individual student). From that point of view tests are meant to rank the students in the ensemble according to their skill and knowledge, while CIs are meant to report the percentage of students in the ensemble that use a particular conceptual construct.

The two descriptions (individual and ensemble) must of course be equivalent, since they are both describing the same underlying system. This is harder than it sounds, and it is the source of most difficulties, both practical and conceptual, in all statistical descriptions of systems from Physics to Economics. What that means, is that for any given case we should come up with the exact same observable answers whether we are looking at the individual view (eg. calculating the likely trajectory of an electron hole in a semiconductor, or the likely portfolio value of an individual investor) or the ensemble view (ie. calculating the total current in the semiconductor, or the total retirement savings of a population). As a practical matter, most fields that use statistical descriptions of their systems have developed more-or-less distinct sub-disciplines that study the two pictures, each with its own distinctive tools and methods. In economics, for instance, the Treasury and the Federal Reserve use macroeconomic tools, theories and measures to follow the economy as a whole, while investment brokers use different tools to produce investment strategies for individuals. The two pictures should be exactly equivalent (and they are rigorously so for systems like ideal gasses, if not necessarily so for the economy) but nevertheless the two sub-disciplines can often look very different.

In education too, different tools and methods have traditionally been associated with individual students than have been used with ensembles. In particular, although tests can be (and sometimes indeed are) used to guide individual students' learning, most tests are mainly used to produce grades (i.e. rankings). Concept Inventories on the other hand are meant to map students' prevalent misconceptions in a field, and hence guide the development of instructional materials and methods that address these misconceptions explicitly. On the student level, CIs can be used to assign supplemental instructional materials that are specifically designed to address that particular student's misconceptions. For example, during the development of the Biology Concept Inventory (BCI), we discovered that an entire class of difficulties that students encounter in both genetics and molecular biology arise from students' misconceptions about random processes (cf. Garvin-Doxas and Klymkowsky, 2008; Klymkowsky and Garvin-Doxas, 2008). In short, students do not understand that processes as diverse as diffusion and evolution are underpinned by random processes which are taking place all the time (molecular collisions and mutations), but think instead that they are driven processes that stop taking place when the driver is removed (they believe that there is no diffusion in the absence of density gradients, and no evolution in the absence of natural selection). This misconception can frustrate learning unless it is directly addressed, and one can envision instructional materials designed to address it explicitly.

As a result of their main use as producers of rankings, tests are therefore (in order of importance) 1) uni-dimensional 2) monotonic, and, as much as possible, 3) linear. Of these properties, the one that mostly defines the structure of a test is linearity.

2.1 Tests as Producers of Rankings

To ensure these properties, test developers look at statistical measures like discrimination (ie. how close can two scores be before we can no longer assure that the higher score indeed represents higher performance) and item difficulty.

Item difficulty is the fundamental weighting factor on which most of the linearization schemes rest. Perhaps the version of difficulty that is most accessible intuitively is the percentage of students that answered the question correctly; questions that have been answered correctly by a large percentage of students have lower difficulty. Item Response Theory (IRT) for instance makes an explicit assumption of true or near unidimensionality, and posits that the probability, P, that a student of ability θ will correctly answer a question of difficulty b is given by the logistic function

$$P = \exp(\theta - b) / [1 + \exp(\theta - b)] \quad (1)$$

Both student ability and item difficulty can then be

place on the same scale (as we see from the logistic function, a student whose ability θ_1 is equal to the difficulty of some item b_1 will have a 50% probability of answering that question correctly).

Difficulty is then used to linearize the response of the test. The most intuitively accessible linearization method, and the one most widely used, consists of constructing the test with questions of many different difficulty levels (b_1-b_4 in Figure-1). The higher the level of the student's skill and knowledge, the more questions s/he will answer correctly. With a large bank of questions to choose from, a test can be devised with questions that are evenly spaced along the difficulty line, effectively calibrating the instrument to insure a more-or-less linear response: answering twice as many questions (above some statistical floor) really means twice the level of performance. We should note here that for multiple choice tests the probability that a student of very low ability answers correctly asymptotes to the random floor (e.g. 25% for a four-option item), but for concept inventories it usually asymptotes well below the random floor, and often close to zero. This is a consequence of having distracters that represent common misconceptions; students who hold an alternative model are lured to the answer that corresponds to their model, and are therefore less likely to pick the correct answer by chance. Statistical treatments that take into account a nonzero asymptote also exist.

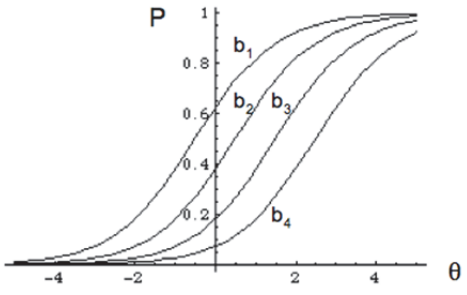


Figure 1: Item Characteristic Curves (ICCs) for four items of difficulty b_1-b_4. Evenly distributing test items along the difficulty line produces a test with linear response.

Recently, more sophisticated linearization techniques like Rasch analysis (Rasch, 1961) have been used for instrument calibration, but all calibration techniques aim for a linear instrument response, and make explicit or implicit assumptions about unidimensionality (or near-unidimensionality). This is a direct and unavoidable consequence of most tests' main use, which is to produce rankings.

2.2 Concept Inventories and Rankings

Necessary as these statistical properties are for tests, they are mostly irrelevant (and sometimes even counterproductive) for Concept Inventories. CIs are by nature multidimensional since what we really want to know is each of the misconceptions that a student holds, not some average over all misconceptions. What we really want to know is what specific instructional material to assign to a student in order to address his/her misconceptions; a measure of the student's average performance level is not at all informative on that task. Furthermore, the percentage of students that answers a question correctly is not an appropriate weighting factor for a CI. The vast majority of the students can, and often do, harbour the same misconception even after repeated instruction; this is the very essence of misconceptions. Leaving these questions out of the instrument, or giving them minimal weight, because they are at the tail of the difficulty distribution is not a productive option.

Nevertheless, CIs have historically been used essentially as tests, reporting a student's improvement in overall performance (i.e. improvement in the total number of items answered correctly) instead of reporting each misconception a student is holding. This use of a CI has proven to be useful in gaining the attention of instructors (e.g. Hake, 1998), and should therefore be considered during instrument constructions as a possible (and even probable) use of the final instrument. That said, results from CIs are inherently much richer in the types of insights they can provide. Given that the objective of a CI is to provide detailed information that can be used to explicitly address student misconceptions, it is useful to have an analysis for each dimension (concept) in the instrument, in addition to an average over all dimensions. This can be done by performing a statistical analysis not only for the correct answer in each question, but also for the answers that correspond to particular misconceptions. In the context of IRT for instance, the Item Characteristic Curve (ICC; Fig.-1) is no longer the probability that the student will answer the question correctly, but the probability that the student will pick the answer that corresponds to a particular misconception, and θ is the student's "ability" *with respect to that misconception*, or in other words the degree to which the student holds that misconception.

The requirement of performing an analysis for each dimension of a CI revives "the curse of dimensionality" (the requirement of analyzing a very

large number of items) which is precisely the problem that modern test theories aim to alleviate. Nevertheless, the requirement is a direct consequence of the function of CIs, which is to produce multidimensional information on the conceptual state of students.

2.3 Validity and Reliability

For an instrument to be useful, be it a test or a CI, it must be valid and reliable. Validity means that the instrument measures what we want it to measure, and doesn't measure things we don't want it to measure (a thermometer should measure only temperature, not some combination of temperature and weight). Reliability means that the instrument gives the same value when measuring identical things. It is obvious from the definitions that validity and reliability are closely related; if an instrument measures only one thing (eg. temperature) then there's only one value it can give (the temperature of whatever we are studying, no matter what its other properties are). It is therefore clear that validity implies reliability. What is less appreciated however, is that reliability does not imply validity. Reliability means that we are consistently measuring the same one thing; but what is that thing? The answer to that question cannot possibly come from the statistics of the instrument alone; an additional input is needed.

That additional input is always theory. The statistics of a reliable thermometer are identical to the statistics of a reliable voltmeter (in fact, most modern thermometers are actually measuring a voltage); the only difference is the theory used to translate the output of the device into a measurement of temperature. In CI construction that additional input is provided by experts who can consistently associate students' verbal cues with persistent mental constructs.

Validation is a labour intensive and time consuming process, the cost of which we can reduce significantly with the use of technology. During the development of the BCI we created Ed's Tools, an online suite of tools that allows us to collect, code, and aggregate large amounts of text data, considerably improving the speed of data collection and analysis. The validation procedure and validation results for the BCI are described in detail in Garvin-Doxas and Klymkowsky, 2008, and Klymkowsky and Garvin-Doxas, 2008. The development and usage of Ed's Tools are described in detail in Garvin-Doxas et. al., 2007. Here we give a short description of the method for completeness, while referring the reader to the previously published work for a detailed exposition.

We start by asking students to provide essay answers to open-ended questions, which we then code using Ed's Tools. The coding allows us to aggregate the language that students use to describe their thinking for each concept that we identify. We then use that language to formulate both the questions and the answers (both the correct answer and the distracters) for the CI items. We then conduct interviews and think-alouds with a large number of students and use these to refine our wording of the Inventory items, and repeat the cycle until the results from the interviews and the instrument converge.

In the following section we describe how we use Latent Semantic Analysis (LSA) to improve the logistics of determining the prevalence of each preconception in the student population, and we show some initial results.

2.4 Latent Semantic Analysis and CI Construction

LSA has been used successfully to provide grading of student essays that correlates well with grades given by experts (Landauer and Dumais, 1997; Landauer et a.l, 1998), and can also be used effectively to provide feedback that helps students (or teachers) identify the elements of the text that they have missed (Kintsch et al., 2000).

In addition to these general language applications, we have recently achieved comparable results in science specific tasks. The results of this work show that with only a small (of the order of ~100) set of human-rated documents to train on, LSA can classify documents that it has not trained on along predefined concept categories in a way that correlates well with the human classification. So far we have analyzed student answers to three questions in Physics, two in Astronomy, and six in Biology.

The Physics results shown in Figure-2 were obtained with data collected with Ed's Tools from three different classes at the University of Northern Colorado (UNC): an introductory calculus-based course for scientists and engineers, and two physics courses for pre-service teachers (an introductory physics course, and a capstone physics course that is required of all graduating pre-service teachers). The essay was assigned during regular class time, and students in all three classes were given 20 min to complete it. The essay was given early in the semester so that the students in the calculus based class and the introductory pre-service class had not covered the material in college.

A typical Physics question was:

In 60 words or more, describe what happens when a light car and a heavy truck, which travel with the same speed but in opposite directions, collide head-on?

As a rule of thumb sixty-word answers are the shortest documents on which LSA can be effective, but with this question we wanted to test LSA's performance for the shortest answers on which the method can be expected to give reasonable results. We collected 65 responses from a class for majors, and a total of 160 responses from two classes for pre-service teachers. Although the overall number of essays we collected was 225, nearly half of them had no physics content (most of the invalid responses concerned seat belt use, insurance rates, and the safety disadvantages of fuel efficient small cars) so the number of relevant essays on which LSA trained was closer to 120. Two expert graders used approximately half of the responses to train on, and scored the remaining half independently. Four rubric components were identified, along which each of the answers was scored on a scale of 0-3. An answer was given a 0 along a component if it did not contain any treatment of the subject, and a 3 if it contained a well articulated treatment (for a misconception, that treatment is physically incorrect, but as long as the concept is clearly present in the text the score for that component is 3). The four components and examples of answers are given in the Appendix. The essays were analyzed using two spaces, TASA, and a physics space constructed for the project. TASA contains 1.2 million words in 37,000 documents and 750,000 sentences and has been selected to be representative of the amount and type of material a college student would have read in their lifetime. The physics space was constructed using

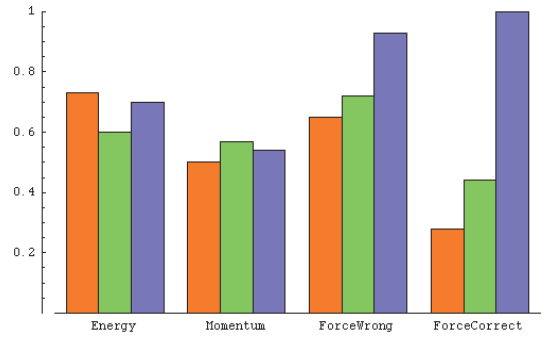


Figure 2: The correlation between the LSA score assignment and the experts' score assignment for each rubric component. The bars represent the TASA-Expert, (TASA+Physics)-Expert, and Expert-Expert correlations respectively for the orange, green, and blue bar.

introductory physics texts available under the Open Content license (http://opencontent.org/opl.shtml) and contains 1465 documents.

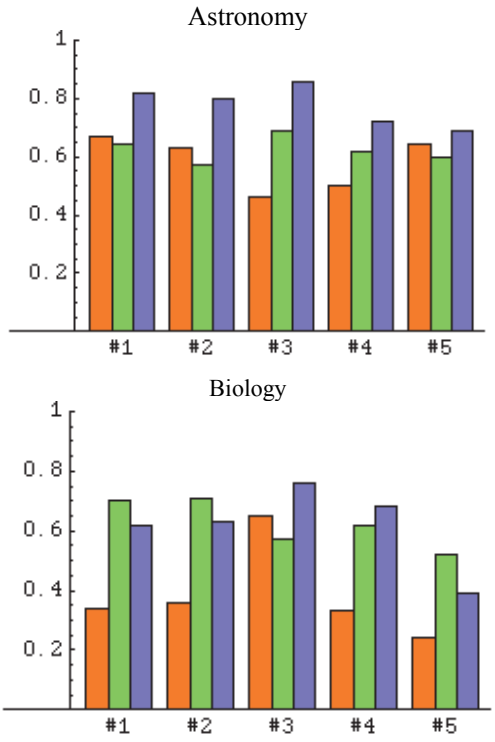


Figure 3: Top frame: The correlation function between the two experts (blue) and between the experts and the LSA system using the TASA general English space (orange) and TASA augmented with the physics space (green). The rubric components are as follows:
#1: The Cosmological Constant (CC) provides a repulsive force that counteracts gravity
#2: The CC is the same as Dark Energy
#3: Study of distant supernovae shows that the expansion of the universe is accelerating
#4: Fluctuations in the microwave background radiation show that the CC exists
#5: Dark Energy is a force that counteracts gravity
Bottom frame: The correlation function between the two experts (blue) and between the experts and the LSA system using the TASA general English space (orange) and TASA augmented with the biology space (green). The rubric components are as follows:
#1: Alternative forms of a gene are known as alleles
#2: Alleles can be dominant or recessive to one another
#3: For most genes, you carry two alleles, one from your mother and the other from your father
#4: A recessive phenotype is visible if both alleles are recessive; if one is dominant, the recessive phenotype will not be visible, but the allele remains and can be passed to offspring
#5: Phenotype refers to the visible traits displayed by an organism.

Figure-1 shows the correlation of the LSA assigned scores (using the two spaces) to the score assigned by expert-1 for each of the components, and the correlation between the two experts. Dimension-3 is the well-known dominant misconception in the domain (that the heavy truck will exert a greater force on the small car than the other way around). We see that LSA is comparable to the experts for component-1 (correct energy formulation) and component-2 (correct momentum formulation), although it performs lower than the experts in component-3 (the dominant misconception on the subject). Component-4 is the correct force formulation of the problem.

It is important to note that TASA alone, which is general space, produces results that are overall comparable to the results produced with the addition of a target-specific physics space. This plot shows that by using human raters to rate a relatively small number of documents, LSA can generally classify documents on which it was not trained, with a correlation which can be comparable to that of different human experts. The exception in this case seems to be the correct force formulation (which states that the forces exerted by the car and truck on each other are equal). It is not clear why this rubric component faired so much worse than the rest. It is worth noting that the experts were in perfect agreement on this component (the correlation is one, over all relevant answers).

Figure-3 shows results from two additional questions, one in Astronomy, analyzed with TASA and the same Physics test used in the Physics questions, and one in Biology, analysed with TASA and an open source Biology text. We see that in both cases the system is consistently comparable to the experts, especially when the general English space is augmented with subject-specific texts.

3 CONCLUSIONS AND FUTURE WORK

Although this is an ongoing project, the results so far show that student essays, even of lengths that are generally on the borderline of being too short for treatment by LSA, can indeed give results that are comparable to expert raters', although some challenges still remain. One of the questions that will be important to the method, is the extend to which the nature of the space in which the texts are projected (eg. a general space like TASA versus a discipline-specific space like the one we developed from the textbooks) affects performance, and we plan to conduct additional studies with a variety of discipline-specific texts to address this question.

Perhaps the greatest limitation of the method is the fact that, at this stage, the dominant misconceptions are still being discovered "by hand" as it were, with experts combing through large amounts of textual data. Tools like Ed's Tools can improve the logistics of that search, and tools like LSA can improve the logistics of identifying these misconceptions in very large populations, but the discovery phase still depends exclusively on experts. We plan to address this limitation in future work, by using LSA to point out possible new misconceptions that can then be rated by content experts.

REFERENCES

Bloom, B. S., The 2 Sigma Problem: the Search for Methods of Group Instruction as Effective as One-on-One Tutoring, *Educ. Res.* 13, 4 (1984).

Garvin-Doxas, K. and M. W. Klymkowsky. Understanding Randomness and its impact on Student Learning: Lessons from the Biology Concept Inventory (BCI). *CBE Life Sci Educ* 7: 227-233 (2008).

Garvin-Doxas, K., I. Doxas, and M.W. Klymkowsky. Ed's Tools: A web-based software toolset for accelerated concept inventory construction. *Proceedings of the National STEM Assessment Conference 2006*. D. Deeds & B. Callen, eds. Pp. 130-139 (2007).

Hake, R., Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *American Journal of Physics, 66,* pp. 64 (1998).

Kalas P, O'Neill A, Pollock C, Birol G., Development of a meiosis concept inventory, CBE Life Sci Educ. 12(4):655-64. doi: 10.1187/cbe.12-10-0174 (2013).

Kintsch, E., D. Steinhart, G. Stahl, Developing summarization skills through the use of LSA-based feedback, *Interactive Learning Environments*, 8 (2000).

Klymkowsky, M.W. and K. Garvin-Doxas. Recognizing Student Misconceptions through Ed's Tool and the Biology Concept Inventory. *PLoS Biology*, 6(1): e3. doi:10.1371/journal.pbio.0060003, (2008).

Landauer, T. K., P. Foltz, and D. Laham, An introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284 (1998).

Landauer, T. K. and Dumais, S. T., A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104(2), 211-240 (1997).

Libarkin, J., and S. Anderson, Development of the Geoscience Concept Inventory, *Proceedings of the National STEM Assessment Conference*, Washington

DC, p. 148-158, (2006).
McDermott, L.C., P.S. Schaffer, and the Physics Education Group at the University of Washington, *Tutorial in Introductory Physics,* Prentice Hall, New York (1998).
Midkiff, K. C., T. A. Litzinger, D. L. Evans, Development of Engineering Thermodynamics Concept Inventory Instrument, 31st ASEE/IEEE Frontiers in Education Conference, Reno, NV (2001).
Rasch, G. On General Laws and the Meaning of Measurement in Psychology. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 4: Contributions to Biology and Problems of Medicine, 321--333, University of California Press, Berkeley, Calif., (1961).
Smith, M. K., W. B. Wood, and J. K. Knight, The Genetics Concept Assessment: A New Concept Inventory for Gauging Student Understanding of Genetics, *CBE Life Sci Educ* 7, 422, (2008).
Zeilik, Michael, C. Schau, N. Mattern, S. Hall, K. W. Teague, and W. Bisard, Conceptual astronomy: A novel approach for teaching postsecondary science courses. *American Journal of Physics 65:10,* 987-996, (1997).

APPENDIX

The rubric components for the Physics example:

Component-1: Energy Conservation
Answers that received a non-zero grade along this component had a correct discussion of energy conservation for the problem. Students usually talked about kinetic energy being converted to other forms of energy during the collision (eg. heat or sound) and correctly stated that the total kinetic energy after the collision is lower than before. Some students even identified and explained elastic and inelastic collisions. The more complete the answer, the higher the score that was assigned to it in this component. For example:

When the light car and heavy truck collide. Each will apply a force to the other. The force from the heavy truck will be greater than the force the car applies to the truck. After the inelastic collision the car will "bounce" off the truck and travel backwards. The truck will slow considerably but should continue forwards. In this collision momentum of the car and truck system will be conserved because momentum is always conserved. Kinetic energy however will be lost because the collision is inelastic. Energy will be lost in the form of heat and sound.

This answer was scored as a 3 in the first component (incidentally, it also scored a 3 in component-3, the dominant misconception in the domain).

Component-2: Momentum Conservation
Answers that received a non-zero score along this component had a correct discussion of momentum considerations for the problem. Students usually talked about the truck having a greater momentum because of its greater mass. They correctly stated that the truck will continue to move in its original direction, while the car will reverse directions, that the combined mass of the car+truck will move at a lower speed than either did before, and many students even stated explicitly that momentum is conserved in the collision. The more complete the answer, the higher the score that was assigned to it along this component. For example:

What happens when the light car and heavy truck collide with each other is that they will have a non-elastic collision. When they crash they will somewhat stick together and continue to move in the same direction as the heavy truck was moving before the collision. The kinetic energy of the light car and heavy truck will not be the same as the kinetic energy of the total mass of the truck and car, because the vehicles are not on a frictionless surface and energy is lost in heat.

This answer scored a 3 in this component (although it is missing an explicit statement for conservation of momentum). It also scored a 3 in component-1 (correct energy treatment) despite the fact that it is ambiguous about the reason for energy non-conservation. Very few answers were better than this.

Component-3: The Force Exerted by the Truck is Bigger
This is the best known misconception treated in the literature. Answers that received a non-zero grade along this component stated that the truck will exert a bigger force on the car than the other way around. For example:

Primarily, when a collision occurs between any object, energy will always be conserved. What will happen in a case where a light car and a heavy truck, traveling at the same speed in opposite directions, collide is each will have a certain magnitude in force and after the collision the vehicles will travel some distance. We know that the heavier truck will have more force because it is more massive. The light car will have less force because it is less massive. The direction in which the vehicles travel post impact depends on the net force resulting between the two vehicles.

This answer scored a 3 in this component. It clearly states the dominant misconception twice,

both for the truck and for the car.

Component-4: Force Equal

This is the correct force formulation for the problem. According to Newton's laws, the force exerted by the car on the truck is equal to the force exerted by the truck on the car. For example:

When a light car and heavy truck collide head on traveling at the same speed the light car will have the most damage. This is not because the force was greater on the car, both are hit with the same amount of force, it is simply because the car is not built as sturdy as the heavy truck.

This answer received a 3 on this component. Some students not only stated this explicitly, but they also quoted Newton's law by name.

The Image of Images as an Aid to Improve Learning
An Eye-tracking Experiment Studying the Effect of Contrasts in Computer-based Learning Material

Mona Holmqvist Olander[1], Eva Wennås Brante[2] and Marcus Nyström[3]
[1]*Department of Pedagogical, Curricular and Professional Studies, University of Gothenburg, Gothenburg, Sweden*
[2]*School of education and environment, Kristianstad University, Kristianstad, Sweden*
[3]*The Faculties of Humanities and Theology, Lund University, Lund, Sweden*
mona.holmqvist@gu.se, eva.wennas-brante@hkr.se, marcus.nystrom@humlab.lu.se

Keywords: Text-picture Integration, Eye-tracking, Dyslexia, Variation Theory, Contrasts in Text and Pictures.

Abstract: This study addresses differences in the design of computer-based learning materials—text with or without pictures—and the aim is to show in what way these differences affect learning outcomes. In total, 46 young adults participated: 19 with dyslexia and 27 controls. Approximately half of each group received the condition *text only*, and half received the text and an integrated picture. The learning material was presented on computer screens, and the participants' viewing patterns were registered by eye-tracking. The respondents answered text-based and picture-based questions, as well as oral questions, during the experiment. The assumptions about learning are based on variation theory and on the importance of contrasts in discerning important aspects of the learning material. The results show that whether material of the same content (surrealism) is presented in text only (without explicit contrasts embedded in the text) or in text and picture form (which offers a contrast) affects learning outcomes, particularly for the participants with dyslexia, who showed a fourfold increase from pre-test to post-test (from .10 to .40).

1 INTRODUCTION

There is an assumption that pictures in learning material support learning, and pictures are frequently used in computer-based learning material. But what happens when learners meet both text and pictures at the same time, as in information graphics? Holsanova et al. (2009) have shown that *how* pictures are integrated with a text is vital for an average reader. When text and pictures are closely integrated, it seems that more of the text is read. Holsanova et al. (2009) measured the number of integrative saccades—that is, "transitions between semantically related pieces of verbal and pictorial information" (Holsanova et al., 2009, p.8)—and thereby found that the quantity of integrative saccades was reduced when the text was separated from the pictures. Furthermore, when pictures and text are serially arranged (as in a classic comic strip) instead of radially arranged (with a large picture in the middle and small pictures and text like satellites arranged around it), the reader tends to read more and for a longer time, and integrative saccades were almost twice as common. One possible explanation for the longer reading time spent on the serially arranged pictures and text was, according to Holsanova et al. (2009), that the pictures and the text were arranged in the direction in which people in Western society read, from left to right. Another is that the text was arranged in a helpful way: it started with general knowledge and built on it with more and more facts. It seems obvious that the way the pictures are exposed, as well as what kind of pictures they are and what they contain, is of importance in terms of where viewers look and for how long. But in what way this contributes to a deeper understanding of the written text has not yet been clarified.

The results presented in this paper originate from a study investigating the way computer-based learning materials with and without pictures affect learning. This was tested in two different groups of subjects, namely, young adults with and without dyslexia. Based on the findings in the main study, further analysis was conducted to describe in what way *contrasts* in the learning material seem to influence the learning outcome more than the

presence or absence of *pictures* in the computer-based learning material does. The aim of the paper is to study how contrasts either within text only or between text and picture affect young adults' abilities to discern and learn from computer-based learning material presented in an experiment about art genres. The study is based on variation theory, and a multi-method approach (Yin, 2006) has been used to measure both the participants' viewing by using eye-tracking and their expressed knowledge by using multiple-choice questions and oral questions during the experiment.

2 THEORETICAL FRAMEWORK

In this paper we take the learning theory approach of variation theory (Holmqvist, Gustavsson and Wernberg, 2008; Holmqvist, 2011; Marton and Booth, 1997; Pang and Ling, 2010; Runesson, 1999) to study which factors are necessary for learning under the two different conditions *text only* versus *text plus picture* in two different groups of participants, young adults with and without dyslexia.

Variation theory supports dual coding theory (DCT; Clark and Paivio, 1991) by assuming that processing printed words and images requires the simultaneous discernment of a part-whole relationship in the object of learning (Svensson, 1976; Gibson and Pick, 2000; Gärdenfors and Johansson, 2005). It also assumes that variety should be designed so as not to overload learners' memory systems (Holmqvist, 2004). Information via the two senses of hearing and seeing does not overload the system as it activates two different working memories (sounds and images), while dual information delivered to one sensory memory can cause an overload (Mayer 2005; 2009). Informed by these studies, our main study (Holmqvist Olander, Nyström, Wennås Brante, manuscript) used only printed words and images—that is, only visual memory—to examine the effect of the part-whole relationship on learning through printed words and images.

2.1 Variation Theory

Variation theory holds that learning takes place through contrast (to make it possible for the learner to separate an aspect from the object), generalization, and fusion. For example, in art genres, which we used in this study, one feature of surrealism is that the paintings are not depictions of reality. Contrasting a surrealistic painting with a pop art painting achieves a contrast regarding the specific aspect of whether the work is reality-based. Through the contrast, it becomes possible to separate this feature from the painting as such. This aspect can be generalized to other art genres, such as impressionism, and it becomes possible for the learner to decide that *Sunrise* by Claude Monet does not belong to the genre of surrealism because the painting does depict reality, even if the expression of it is a bit hazy. Such contrasts can be offered to the learner in various ways; in our case, we chose to present the contrasts either in text or in a combination of pictures and text so as to study whether and how the different types of material facilitate or complicate learning.

2.2 Variation Theory and Reading

To read is to combine signs with sounds, putting sounds together into meaningful words and interpreting them in the context in which they are found. Drawing on variation theory about how knowledge is attained (Marton and Tsui, 2004), a person who is reading needs to simultaneously experience both the parts—that is, the phonemes represented by one or more letters—and the wholeness—that is, the words in that particular text as well as the entire text itself. One example (Wennås Brante, 2010) involves the difficulties involved in differentiating *b* from *d*. If a child sees a dog, it does not matter whether the dog turns around; it remains the same dog. But the signs *c* and *l* are the parts that make up wholes as they can be used both independently and in any combination. The combinations can be *b*, *d*, *p* and *q*. The critical feature is to know the importance of the parts' directions in the combination so as to understand that the letters differ and not are the same when seen from another perspective. To make the distinction, learners must discern the parts simultaneously with their directions. If they learn only about *b*, there is a risk that a *d* or *p* will be interpreted as a *b* the first times the learners see the new letters. If they learn about *b*, *d*, *p*, and *q* at the same time, the variation needed to discern the importance of the parts' directions will be offered. Based on this example, it is easy to understand the complexity of the critical features that reading comprehension includes. Through the use of interviews and eye-tracking it is possible to ask participants what they have actually been aware of and can express about the text, at the same time as we study what they *could* have been aware of by following their eye gazes. This differs from what they have been *presented*, as we do not

know which parts of the text they seem to neglect and which parts they seem to focus on.

2.3 Picture Viewing

Perceived objects are affected by visual factors as contrasting or semantic factors (Henderson, 2003; Nyström and Holmqvist, 2008). This means that people use their knowledge about the world to gaze intelligently through a real-world scene (Henderson, 2003). They seem to do the same thing when looking at a picture. Yarbus (1967) showed in a famous eye-tracking study that a task influences the visual pattern. When the task was to find out relationships between people in a picture, one kind of a scan-path emerged which differed from the scan-path emerging when the task was to find out whether, for example, the people in the same picture have eaten dinner. Viewers tend to know what they are looking for and steer their fixations in that direction, according to Yarbus (1967).

Brockmole and Boot (2009) have reported that task-irrelevant aspects of visual stimuli also influence the fixation durations. This complicates the use of pictures as comprehension aids because it is hard to ensure that every picture tied to a text has only task-relevant aspects. If it does not and if the reader encounters problems comprehendingthe parts simultaneously with the whole in a single gaze, the pictures become just one more part to be distracted by. Furthermore, a study by Nummenmaa et al. (2006) shows that people seem to allocate attention to emotional pictures rather than to neutral ones. Even when instructed to attend to neutral pictures, fixations on the emotional ones form. This means, according to Nummenmaa et al. (2006), that orientation towards emotional stimuli is highly exogenous. One cannot, in other words, completely steer one's attention to the pictures that are perhaps most relevant or that contain the most information; one is drawn to emotional pictures. Text and picture integration can thereby be efficient or misleading, depending on which pictures are presented.

A study (Hannus and Hyönä, 1999) of 10-year-old children showed that high-ability children paid more attention to pictures in a text and could make better use of them than low-ability children did and could. Comprehension scores were also higher for high-ability children when they encountered a text with illustrations, but comprehension scores for low-ability children did not improve under the same conditions (Hannus and Hyönä, 1999).

Holsanova et al. (2009) showed that *how* pictures are integrated with a text in information graphics is vital for an average reader. The way pictures are designed and offered, as well as what they contain, is important in terms of where a viewer looks and for how long. However, this does not tell anything about the viewer's awareness. How the viewing pattern contributes (or does not) to a deeper understanding of the written text has not yet been clarified.

A person reading a text processes it visually, phonologically, and semantically at the same time. When reading a text integrated with pictures, the person attends to the pictures, as well. What does this mean for dyslexic readers' reading comprehension?

3 PREVIOUS FINDINGS

The results of the main study show that there are significant differences between the two groups of participants and interesting differences between the two different conditions. The results show significant differences in both comprehension scores and visual behaviours between respondents with and without dyslexia in the two different conditions (text only or text plus picture). Respondents with dyslexia did not use strategies to gain a global overview of the information, something that the participants without dyslexia did more quickly. This trend supports the hypothesis that the participants with highest learning outcomes used global-overview strategies. Participants with dyslexia showed haphazard viewing patterns, strengthening the assumption that a global overview as a strategy promotes learning (Holmqvist Olander, Nyström & Wennås Brante, manuscript). The condition *text and picture* decreased learning for respondents with dyslexia, with one exception: the example of surrealism, which this paper examines further.

3.1 Dyslexia and Phonological Awareness

One pre-requisite to reading is discerning a given sound among the continuous flow of sounds people produce when they speak (Shaywitz, Morris and Shaywitz, 2008) and then mapping the sound to a letter (Byrne, 2005). This phonological awareness is one part of the phonological ability needed to read; thus, a weak phonological ability is considered one explanatory factor in dyslexia (Goswami and Bryant, 1990; Ramus et al., 2003; Vellutino et al., 2004; Everatt and Reid, 2009). In the main study, participants were tested for phonological ability

using a screening test called Duvan (Lundberg and Wolff, 2003). The test is executed in groups and tests for different sub-skills in phonological ability. The results from the main study (Holmqvist Olander, Nyström, Wennås Brante, manuscript) show a significant difference between the groups, supporting the assumption that participants with dyslexia have a lower phonological ability than the rest of the participants do.

3.2 Viewing Patterns

In the experiment, pictures were inspected for less time than texts were, yet a difference between groups was found. The control group performed an earlier inspection of the picture, thus rendering a global impression of the material (Holmqvist Olander, Nyström, Wennås Brante, manuscript). Viewing analysed only in the condition of text + picture and the picture was viewed by 95.2% of the controls and 78.3% of those from the participants with dyslexia. The participants spent 45.9 seconds (*SD* = 2.1) looking at the text and 2.2 seconds (*SD* = 2.3) viewing the picture. They were looking elsewhere for 0.8 seconds (*SD* = 1.0). When interviewed about viewing patterns, some respondents with dyslexia had an understanding of the way they processed material containing both text and pictures that was found inconsistent with how they really placed their gazes (Wennås Brante, 2013). They processed the material in a much more random way then they thought they did, thus missing opportunities to gain a global understanding of the material.

3.3 Contrasts

The texts from the experiment in the main study contained a contrast formulated in words. The contrast could be explicit (that is, the two values were specified and mentioned) or implicit (that is, one value not was specified). The parts of the texts containing a contrast were the ones most discerned (Wennås Brante, Holmqvist Olander and Nyström, 2013), even when the contrasts were of less importance for learning about the content. If a contrast was implicit—that is, if a counter-value was not specified—it was hard for the dyslexic group to discern it (Wennås Brante, Holmqvist Olander & Nyström, 2013). The micro-analysis of the oral answers revealed also that contrasts are exceedingly effective for discernment (Wennås Brante, Holmqvist Olander and Nyström, 2013).

The experiment also employed different kinds of pictures that could influence the viewing behaviour. Three pictures differing in motif and color were chosen to test how the viewing behavior differed: a cubistic picture that had a motif that was difficult to see in brown nuances, a romantic picture with a clear motif painted in natural colors, and finally a surrealistic picture with an odd motif and clear colors. No great difference was found in viewing patterns of the cubistic and romantic picture, but participants without dyslexia had significantly higher dwell times on the three pictures ($p = 0.034$), a result that originated from dwell time on the surrealistic picture (Wennås Brante, manuscript). However, most participants with dyslexia inspected the surrealistic picture before they started reading the text, and this had an impact on their reading comprehension: the oral answers from the group with dyslexia contained more inferences regarding the surrealistic image than the same group's answers about the other art genres did (Wennås Brante, manuscript).

4 IMPLEMENTATION

The analysis is based on data collected during the main project (Holmqvist Olander, Nyström and Wennås Brante, manuscript; Wennås Brante, Holmqvist Olander and Nyström, 2013, Wennås Brante, 2013; Wennås Brante, manuscript). So far no analysis or results have been presented on the task level regarding the outcomes of the questions in the experiment.

4.1 Participants and Design

In total, 46 respondents were recruited via an advertisement on two universities' websites and in a residential college for adult education through a teacher, requesting young adult native speakers of Swedish with and without dyslexia and with normal or corrected-to-normal vision. In total, 19 dyslexic and 27 non-dyslexic (control) respondents were recruited. Their ages ranged from 19 to 41, with a mean age of 23.8 years (*SD* = 4.2).

The reading tasks consisted of six different short expository texts which demand somewhat different reading strategies than narrative texts do (Hyönä et al., 2002), such as understanding the topic and the hierarchically related sets of sub-topics and their relationships with one another. Because expository texts are frequently used in schools and education, it is of interest to determine whether pictures contribute to the comprehension of expository texts.

Respondents were presented first with the name of a genre, then with information from either of the two experimental conditions (*text only* or *text plus picture*). The computer-based learning material in the two conditions was identical, except from text only in one condition and text with pictures in the other condition. In total it consisted of six screens; 1.Title of art genre, 2. Text about the art genre (including picture in one condition and without in the other), 3. Blank screen (the researcher asks what information the respondent met on the previous screen – oral post-test), 4. Text question (verbal post-test), 5. Confidence rating and 6. Multiple-choice question illustrated by three paintings (picture post-test). When the respondents had studied the information at the second screen and were ready to answer questions about it, they clicked the mouse, and an empty screen appeared. While instructed to keep looking towards the screen area, respondents were by one researcher asked "What information about [genre] appeared on the last screen?" Their verbal answers were recorded through a webcam. When they had no more information to provide, they were asked to click the mouse and proceed to answer three multiple-choice questions with only one of the alternative answers. The first question offered a choice of five text responses to a question about the genre. In the second question, respondents were asked to estimate how confident they were about their previous answer on a scale of one (very uncertain) to seven (very certain). Finally, they answered the same questions again as they had during the initial test, with the positions of the three pictures randomly shuffled, in order to assess whether learning had taken place during the experiment (Holmqvist Olander, Nyström and Wennås Brante, manuscript).

The eye-tracker used in the main study was an SMI HiSpeed sampling data binocularly at 500 Hz. This speed is sufficient to measure saccades accurately. Stimuli were shown on a 19-inch screen connected to a PC which also collected each subject's eye-tracking. To ensure that it would be possible to detect which work a person was looking at, text was presented with letters spanning approximately half a degree of visual angle. Rows had double line spacing.

5 RESULTS

The results of the main study show significant differences among the two groups of participants regarding both viewing patterns and learning outcomes (Holmqvist Olander, Nyström, Wennås Brante, manuscript; Wennås Brante, Holmqvist Olander, Nyström, 2013; Wennås Brante, 2013, Wennås Brante, manuscript). However, analysing the computer based learning material at the task level shows one exception regarding an item in which both groups exhibited the same kind of strategy—namely, the task regarding surrealism. This finding was first measured by the eye-tracking equipment (Wennås Brante, manuscript). The results of the analysis indicate that all participants used the same kind of viewing pattern when viewing the painting of surrealism. The analysis of the participants' answers on the text and picture questions supports this initial indication and demonstrates a correlation between the pattern of eye movements and learning outcome. The result contradicts previous findings as it highlights the impact of the design of the computer-based learning material regarding illustrations. Pictures or not is not a simple question that can be stated to improve or worsen the learning outcome. However, the text material about surrealism did not offer contrasts that are significantly pointed out, resulting in fewer discerned contrasts in the task on surrealism (20–64%) than in, for instance, the task on impressionism (65–77%). Further, this difference between the groups is significant (dyslexic 20–33%, controls 41–64%) for the surrealism task but not for the impressionism task (dyslexic 65–72%, controls 65–77%) (Wennås Brante, Holmqvist Olander, Nyström, 2013). This indicates that contrasts that are not clearly expressed are harder to discern for those with dyslexia than they are for the rest.

In the condition *text plus picture*, the surrealistic picture itself comprises a contrast with reality as experienced by the participants; moreover, the picture is viewed to a higher degree than the other pictures (Wennås Brante, manuscript). This is true also for the group of dyslexics, and they do not show any similarities with the controls' viewing pattern regarding the other stimuli. The result shows that the same learning content presented by two different representations, text only or text and picture integrated, is perceived differently as a consequence of the discerned contrasts offered by the computer-based learning material. This is shown by the increased learning outcome regarding the condition *text plus picture*, as the participants' discernment of the surrealistic painting contrasted with their own views of reality when they observed the aberrant painting. It is more difficult to draw such a strong contrast when present with the text-only stimulus about surrealism. The results of the pre- and

post-tests from the picture question appear in table 1.

The text in the learning material about surrealism describes objects in a way that differs from their ordinary use in reality (e.g., 'Instead, a connection between, for example, a green hot dog and a gigantic comb wer necessary'). But not until participants see the paintings can they fully grasp how different the image is from what is possible in reality. If the results of the participants' answers from the text-based questions and the picture-based questions about surrealism are broken out from the material previously presented, a new and different pattern emerges in the group of dyslexics regarding this specific genre in relation to the mean of the remaining five genres.

Table 1: Means of the results from picture-based question about surrealism in comparison to the means of the other five tasks in the two different groups of participants.

Mean points/ question	Dyslexic Text only	Non-dyslexic Text only	Dyslexic Text plus picture	Non-dyslexic Text plus picture
Pre-test Surrealism	.44	.38	**.10**	.21
Post-test Surrealism	.78	.77	**.40**	.50
Pre-test Other genres	.24	.37	.26	.36
Post-test Other genres	.60	.66	.68	.63

Only three participants in the entire group failed to correctly answer the picture question related to surrealism. Of these three, one came from the group in which the participants did not see the picture (*text only*). The other two were in the group with the condition *text and picture*, but one of those had not fixated at all on the picture (the only respondent to demonstrate that behavior), and the other one had fixated on the picture only once.

The results in table 1 show that there is a learning outcome for the group of participants with dyslexia in the condition *text plus picture* regarding the task on surrealism that differs from the ordinary pattern—a fourfold increase—more than for any other genre or condition. In the condition *text only*, there is no difference in outcomes between the questions for surrealism and the questions on the other art genres; the participants scored roughly twice as high on the post-tests as on the pre-tests.

Regarding the text-based question about surrealism, the group of participants with dyslexia showed slightly higher scores (.80) for the task about surrealism than the mean scores for the other genres (.68), but this is far from the fourfold increase apparent for the picture-based question (table 2). This supports the assumption about contrast's impact on learning: there are implicit contrasts in the text, resulting in lower scores for the *text only* condition for both text- and picture-based questions. On the other hand, there was a high increase for the condition *text plus picture* regarding the picture-based question about surrealism.

Table 2: Means of the results from the text-based question about surrealism compared to the mean of the other five tasks in the two different groups of participants.

Mean points	Dyslexic Text only	Non dyslexic Text only	Dyslexic Text plus picture	Non-dyslexic Text plus picture
Surrealism	.89	1.00	.80	1.00
Other genres	.89	.91	.68	.86

6 CONCLUSIONS

The main finding in this study is the importance of purposeful designing computer-based learning materials in relation to content and learner; in particular, it is expedient to offer contrasts embedded in text or pictures in order to make it possible for the viewer to grasp the meaning of the information. The results show that participants with dyslexia are particularly helped by explicit contrasts—for instance, demonstrating what something is by also showing what it is not—and their learning outcomes are also affected to a higher degree than those of individuals without dyslexia. Illustrations in computer-based learning material have to be meaningful and purposeful designed, and not be used as decorations, to influence learning. Findings about decreased learning, when using learning material with illustrations for students with dyslexia, point out the need of further research about how illustrations carefully can be used to improve instead of prevent learning. The results from the eye-tracking data validate the differences in viewing-pattern in between the groups of participants and reveals the strategies used when examining computer-based learning material.

ACKNOWLEDGEMENTS

This project is funded by the Swedish Research Council (project number 2010-5379), for which we

are grateful. The project is also supported by University of Gothenburg, Lund University, and Kristianstad University. We thank all the participants for their contribution to the study, as well as professor emeritus Ference Marton, University of Gothenburg, for valuable comments through the entire project.

REFERENCES

Brockmole, J. R., Boot, W. R., 2009. Should I stay or should I go? Attentional disengagement from visually unique and unexpected items at fixation, *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 808-815.

Byrne, B., 2005. Theories of learning to read. In M.J. Snowling & C. Hulme (Eds.) *The science of reading: A handbook*, 104–119. Blackwell publishing. Malden, Ma.

Clark, J. M., Paivio, A., 1991. Dual coding theory and education, *Educational psychology review*, *3*(3), 149-210.

Everatt, J., Reid, G, 2009. Dyslexia: An overview of recent research. In G. Reid (Ed.) *The Routledge Companion to Dyslexia,* (pp. 3-21). Routledge. London.

Gibson, E. J., Pick, A. D, 2000. *An ecological approach to perceptual learning and development*, Oxford University Press. USA.

Gardenfors, P., Johansson, P. (Eds.), 2005. *Cognition, education, and communication technology,* L. Erlbaum Associates. Mahwah, N.J.

Goswami, U., Bryant, P, 1990. Phonological skills and learning to read, *Journal of Child Psychology and Psychiatry*, *32* (7), 1173-1176.

Hannus, M., Hyona, J., 1999. Utilization of illustrations during learning of science textbook passages among low-and high-ability children, *Contemporary Educational Psychology*, *24*(2), 95-123.

Henderson, J. M., 2003. Human gaze control during real-world scene perception, *Trends in Cognitive Sciences*, *7*(11), 498-504.

Holmqvist, M., 2004. *En främmande värld. Om lärande och autism.* [A Strange World. About Learning and Autism.] Lund: Studentlitteratur.

Holmqvist, M., Gustavsson, L. & Wernberg, A., 2008 Variation theory – An organizing principle to guide design research in education. In Kelly, A.E., Lesh, R., &. Baek J. (Eds.) *Handbook of design research methods in education, p 111–130*. New York: Routledge.

Holmqvist, M. 2011. Teachers' learning in a learning study. *Instructional Science, 39*(4), 497–511.

Holmqvist Olander, M., Nyström, M., Wennås Brante, E., manuscript. *Eye-tracking as an indicator of impaired executive function in people with dyslexia.*

Holsanova, J., Holmberg, N., Holmqvist, K., 2009. Reading information graphics: The role of spatial contiguity and dual attentional guidance, *Applied Cognitive Psychology*, *23*(9), 1215-1226.

Hyönä, J., Lorch Jr, R. F., Kaakinen, J. K., 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns, *Journal of Educational Psychology*, *94*(1), 44-55.

Lundberg, I., Wolff, U., 2003. *DUVAN. Dyslexiscreening för ungdomar och vuxna,* Psykologiförlaget. Stockholm.

Marton, F., Booth, S., 1997. *Learning and awareness*, Lawrence Erlbaum. Mahwah, N.J.

Marton, F., Tsui, A., 2004. *Classroom discourse and the space of learning,* Lawrence Erlbaum Associates. Mahwah, NJ.

Mayer, R. E., 2005. Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning,* pp. 31-48, Cambridge University Press. Cambridge.

Mayer, R., 2009. Multimedia principle. In R. Mayer (Ed.), *Multimedia learning* 2nd ed, pp. 223-241, Cambridge University Press. New York.

Nummenmaa, L., Hyönä, J., Calvo, M. G., 2006. Eye movement assessment of selective attentional capture by emotional pictures, *Emotion*, *6*(2), 257-268.

Nyström, M., Holmqvist, K., 2008. Semantic override of low-level features in image viewing-both initially and overall, *Journal of Eye-Movement Research*, *2*(2), 2-1.

Pang, M. F., Lo, L. M., 2012. Learning study: helping teachers to use theory, develop professionally, and produce new knowledge to be shared. *Instructional Science*, 1-18.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., Frith, U., 2003. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults, *Brain*, *126*(4), 841-865.

Runesson, U., 1999. *Variationens pedagogik. Skilda sätt att behandla ett matematiskt innehhåll.* [The pedagogy of variation: different ways of handling a mathematical topic], Acta Universitatis Gotheoburgensis. Göteborg, Sweden.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. Annu. Rev. Psychol., 59, 451-475.

Svensson, L., 1976. *Study skill and learning*, Universitatis Gothoburgensis. Göteborg, Sweden.

Vellutino, F., Fletcher, J., Snowling, M., Scanlon, D, 2004. Specific reading disability (dyslexia): What have we learned in the past four decades?, *Journal of child psychology and psychiatry, 45*(1), 2-40.

Wennås Brante, E., 2010. *Identifying critical aspects from learners' perspective*. Paper presented at The 8th International Conference on Education and Information Systems, Technologies and Applications: EISTA 2010, Orlando, Florida, USA, 29 June – 2 July, 2010.

Wennås Brante, E. (2013). Att orientera sig i text och bild. Skillnader mellan förmodad och faktisk läsning för läsare med dyslexi. Acta Didactica Norge, 7(1), (Art. 12, 20 sider).

Wennås Brante, E. (manuscript). Non-congruent pictures

effects' on reading comprehension: *Differences and similarities in patterns of transition between text and picture among young adults with and without dyslexia*.

Wennås Brante, E., Holmqvist Olander, M., Nyström, M., 2013. Exploring the impact of contrasting cases in text and picture processing*, Journal of Visual Literacy*, *32*(2):15–38.

Yarbus, A. L., 1967. Eye movements during perception of complex objects, *Eye movements and vision*, *7*, 171–196.

Yin, R. K., 2006. Mixed methods research: Are the methods genuinely integrated or merely parallel. *Research in the Schools*, *13*(1), 41–47.

Mediator-learner Dyad
Cooperative Relationships

Gilda Helena Bernardino de Campos[1,2] and Gianna Oliveira Bogossian Roque[2]
[1]*Dept. Education and Central Coordination for Distance learning, Pontificia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 22453900, Brazil*
[2]*Central Coordination for Distance learning, Pontificia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 22453900, Brazil*
{gilda, gianna}@ccead.puc-rio.br

Keywords: Pedagogic Mediation, Distance Learning, Teacher Performance, Alceste.

Abstract: This study presents results from a quali-quantitative research looking for evidence on the pedagogical mediation influence on the academic performance of distance learning students. For the systematization of the responses to the research question, a content analysis was performed using the Alceste method. An explanation for the evidence of the mediation relevance to the student professional performance, following from the dyad for cooperation concept, based on social psychology theoretical underpinnings, founded on the philosophy of the Hegelian logic. These concepts paved the ground on which the analysis was deployed.

1 INTRODUCTION

The accelerated increase in the number of undergraduate and post-graduate courses delivered through the distance education model in Brazil in recent years makes it necessary to rethink the indicators used in the quality assessment of distance learning courses.

This article presents the results of a research on the actions of pedagogical mediators. We assumed that the relationship established between mediator and student forms a dyad that establishes a situation of cooperation throughout the course, as defined by Morton Deutsch (1973). This means that there is only one goal for these participants connected in such a way that the final goal can only be reached if, and only if, both are able to reach their goals. Deutsch coined the term *promotive interdependence* which has been used to characterize the relationships and connections in which there is a positive correlation between achievement and success of both participants (p. 20). From a philosophical viewpoint, we believe that this positive correlation, as well as its counterpart, the negative correlation of competition, taken as opposites, can be expressed in the contradiction of Hegelian dialectic (Hegel, 1812).

2 THE FIELD OF RESEARCH

The central objective of the present research was to investigate the way students signified what is quality in distance learning courses. To this end, we developed a questionnaire consisting of 25 objective questions and four open-ended questions – made available online and protecting the identity of respondents. This instrument was sent to all 3,828 students finishing the specialisation course Tecnologias am Educação [Technologies in Education], resulting in 2,117 answers.

The specialisation course Tecnologias em Educação targeted teachers and managers selected by the State Departments of Education (SEE) and Municipal Departments of Education (SME) covering the entire Brazilian territory in different regions.

The structure of monitoring and support of the course participants involved teachers-authors, pedagogical mediators, educators on different nuclei of educational technology in the country, coordinators, technical support, secretariat, among others. In order to academically monitor course participants, distributed in 208 classes, a team of 110 pedagogical mediators was formed, and each was responsible for one or two classes of a maximum of 30 students. It is worth noting that the course was distance-learning, and the pedagogical mediator was

responsible for the enhancement of the learning space, individualised attention to students, guidance with regard to the content covered in the disciplines by establishing a dialogue between students and the study materials and learning assessment.

The objective questions of the questionnaire sought evidence on the following topics: teaching material; study guidance; student support; technical support; study environment, and self-assessment. For writing this article we selected part of the study which deals with one of the open-ended questions on pedagogical mediation: *How did the Pedagogical Mediation carried out throughout the course contribute to your academic performance?* The procedure we chose for data analysis was discourse analysis.

3 PROCEDURES FOR DATA ANALYSIS

In this research, we chose to work with content analysis through the ALCESTE software: Analyse des Lexèmes Coocurrents dans un Ensemble de Segments de Texte (Reinert, 2007, p. 35). Either in the analysis of co-occurring lexemes in a set of text segments or in simple statements in a text, "the method's principle is simple," writes Reinert (2007): "The corpus to be analysed is cut into a series of text segments and the distribution of full words in these segments is observed: hence the name of the method" (p. 35). The "full word," assumed as something meaningful, is different from the "tool word" ("mot outil") like, for example, the conjunction or pronoun that acts as an aid supporting language in discourse.

Its functionality consists in successive analyses that seek to distribute the co-occurrence of the full word in different text segments or in its simple statements, taken two by two. Within the theoretical framework which conceived and operationalised the Alceste method, we find the fundamental assumptions of psychoanalytic theory on the concepts of free association and repetition, from which the method has been gradually implemented.

The conception of the ALCESTE method leads Reinert to Lacan and demarcates precisely what one should expect from key-algorithms that intend to seek significant structures, that is, configurations signalling meaning:

> Studying the meaning of an utterance, of a discourse, presupposes that one takes into account a specificity of meaning. It is trinitarian: meaning as sensation, meaning as direction, and meaning as significance. The meaning of an utterance develops dynamically due to a reading like a braid with three strands […] the analysis of certain corpus allows us to undo the knots of the braid (REINERT, 2000, p. 1).

Reinert did this in order to bring these concepts to the data table to be prioritised, without any kind of a priori criteria, through correspondence analysis.

In practical terms, the Alceste software seeks a pragmatic approach to the text, focused on lexical co-occurrence, co-presence of vocabulary contrasted to two contextual units of the corpus. According to Lima (2008), it is not about comparing the statistical distribution of 'words' in different data corpora, it consists in studying the formal structure of their co-occurrences in the 'statements' of a given corpus. What the Alceste programme examines in the text are not the superficial divisions resulting from the meanings of individual words, but, as Lima (2008, p. 89) points out, "it examines the resonances of meanings established due to the co-occurrences of some terms, and which appear clumped together in certain parts of the text (discourse) at certain points." Later, we may observe the analysis stages or procedures of the software.

To enable the organisation of the analysis procedures we first needed to delimit the categorisation of the text to be analysed, the occurrence of which is registered according to the categories raised and the contextual units – the broadest part of the content to be analysed. We used as variables some characteristics we regarded as significant to the analysis of the questions, in this case: the number of the individual, their gender, the class to which they belong, and their school system.

4 DESCRIPTION OF QUALITATIVE DATA

We identified 2,079 units of initial context (UIC), that originated, from the analytical approach of the Alceste software, 2,049 units of elementary context (UEC). Of those, 85% were classified, resulting in the hierarchisation of four classes.

From the data matrices, the program performs a descending hierarchical classification (DHC), which produces a hierarchical tree, the dendogram of Figure 1, which reveals the identification of the classes through the DHC, where we can identify two groups. The first, formed by Class 3, and the second formed by classes 2, 4, and 1. Camargo (2005) explains that the descending hierarchical classification only ends when the classes become stable. In the specific context of the linguistic

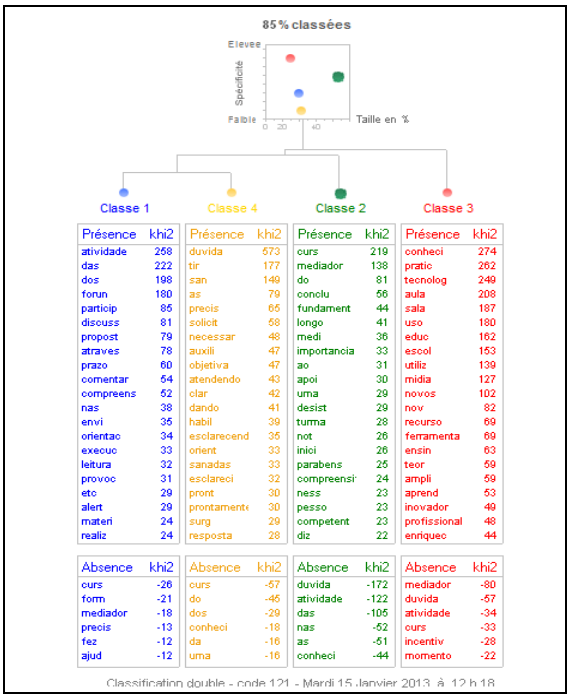


Figure 1: Descending Hierarchical Dendogram.

intersection with psychoanalytic theory (Reinert, 2001) there is an observation we consider important in the assessment of the stabilisation of the process of hierarchical descent created by Reinert for the Alceste methodology, technically enabled by Benzécri's (1992) correspondence analysis.

We note that Class 3 is the class with the highest number of occurrences, therefore, it is not the biggest in size. Class 2, in spite of having the highest number of occurrences, is the second in hierarchy and contrasts with Class 3. This contrast contains the central point of this article, which consists in the dyad between mediator and student.

The goal of the descending hierarchical classification is to make explicit "among all possible partitions in two classes the one which can be considered the most significant statistically" (Marchand, 2007, p. 63). It is an iterative procedure – the first class analysed comprises all context units retained, and the procedure seeks the partition in two classes which makes the chi-squared of the table margins reach their maximum. The most "important" and significant class comes from there. The same procedure is carried out with what remained outside this class and which will again produce two classes, one of which will be retained, the rest reprocessed in new iterations, until the procedure is exhausted and stops.

Within this technical framework, let us comment on some of the main results. Class 2 is subdivided into two classes, forming a perfect hierarchy. It is a "bipolar hierarchy," to use the terminology used by Kalampalikis (2003) for this descending sequence configuration. This hierarchy points towards the existence of the contradiction, in the dialectical sense proposed by Hegel, between two opposed individuals who have a cooperative position but who belong to a different locus from one another and pursue the same goal. In this point Deutsch's (1973) theory of cooperation and competition comes in, based on the contradiction of opposites in a situation of interdependence, operationalised in the same two-by-two matrix we have been using. In the cooperative situation, the correlation is positive in the intertwining of the goals achieved by participants. In the competitive situation, in turn, the correlation is negative, because the intertwining of goals achieved is based on contradiction. The application of these concepts of cooperation/competition in the educational context, within the framework of the theory of social interdependence, is exemplified by the great impact of the work, in *Educational Researcher*, by Johnson (2012).

The divisions between classes represent the relation of proximity of pre-existing meanings common to them. Considering the reduced words and/or forms, it is possible to name the formation of classes in the following way:

- Class 3 – pedagogical training of mediator;
- Class 4 – Mediator as the main motivator of the student;
- Class 1 – Pedagogical strategies of mediation;
- Class 2 – Mediator as a "knowledgebase."

4.1 Analysis of Class 3 – Pedagogic Training and Mediation

Class 3 deals with essential aspects to the training of the pedagogical mediator from the viewpoint of students in the professional training of the participating teachers. In this way, we managed to identify two distinct axes in one single class. The first axis includes the words slogans learn+, resource+, teach+, educ+, media+, use+, class+, and room+ [*aprendi+, recurso+, ensin+, educ+, midia+, uso+, aula+*, and *sala+*] as those related to the pedagogical mediator's theoretical and practical knowledge. The second axis, in turn, includes the radicals professional+, enrich+, knowledg+, ampli+, reflect+, practic+, and innovat+ [*profissional+, enriquec+, conheci+, ampli+, reflet+, pratic+*, and *inovador+*], that are revealing of the result or

influence that pedagogical mediation had in the professional performance of the student.

In this class, two aspects are highlighted: Educational background of pedagogical mediator and Influences of mediation on the professional practice/training of the student. In the students's conception, the pedagogical mediator, in their training, must have knowledge on the learning processes, didactic in distance education, and theoretical and practical knowledge on the subject they teach. The knowledge of the pedagogical mediator in relation to learning processes are intimately connected to the pedagogical and communicational conception on which the course is based. Within the skills that involve learning processes, the students raised questions about the commitment the mediator should have to student learning.

> Pedagogical mediation is the basis of the process of teaching and learning and in a distance learning course, and essential for the permanence of the student and construction of knowledge (Ind_344 *Gen_F *Turma_BA05)

With respect to didactic in distance education, the students presented two aspects in their statements,: the first, that the mediator needs to know distance education how to use the interfaces present in the virtual environment, and teach students how to work with the environment's interface, and enable student-mediator interaction in the virtual environment. Another aspect to be highlighted is the fact that many students had their first contact with distance education and did not feel confident in the navigation. It was up to the pedagogical mediator to identify these students and offer them support.

> (…) it was very important, because I managed to learn how to interact and handle the several tools and new technological advances, besides entering the era of digital inclusion, because I was ignorant on the subject (*Ind_1092 *Gen_M *Turma_PB01).

Here we emphasize the importance of a mediator who is a researcher, in the sense of always seeking new media and new uses to share with students, as well as encouraging them to develop an autonomous way of learning, carrying out their own researches according to their school routine.

The influence of mediation on the practical and professional education portrayed by the students was understood as reflection and performance in classroom. The recognition of reflection as something meaningful in the professional performance of participants can be perceived in the following statements:

> Pedagogical mediation was precise, providing content that brought reflection on the pedagogical performance of each student in their working environment, reevaluating practice (*Ind_1590 *Gen_F *Turma_RO03).

From the student's statements we can observe that the influence of reflection was not attributed only to the pedagogical mediator, but also as a result of the whole course, being seen as a skill that was developed be means of learning, collaboration, and knowledge built during the specialization.

It is through the act of reflection that the teacher feels the need to change their pedagogical practice or, as was brought up by the students, their performance in the classroom. It is not enough for the teacher to only know how to work with technologies, manipulate them, or include them in the school routine and in the class plan if this process is not accompanied by reflection. Their performance in the classroom will not significantly change.

> Pedagogical mediation provided some didactic innovations that facilitated my teaching practice. (*Ind_1957 *Gen_F *Turma_SP04)

4.2 Analysis of Class 2 – Motivation Strategies

Class 2 is hierarchically superior to Classes 1 and 4 and contains 44% of the analysed UECs. We identified in this class UECs related to the mediator as someone who motivates the student during the course and about the assessment strategies used by the mediator. We also found statements from students who attributed to the mediator their having been able to complete the course.

We note a division between two larger groups, with the first one also divided (Classes 1 and 4). The meanings present in the first group (Class 2) are: the importance of the mediator in the conclusion of the course, based on the radicals mediator+, essent+, importance+ [*mediador+, fundament+, importancia+*]; and the characteristics or attributions of the mediator in order for them to motivate the student, which we identified through the radicals competent+, sympathetic+, drop+, congratulations+ [*competent+*, *compreensiva+*, *desist+*, *parabéns+*]; and support from the mediator for the student to conclude the specialisation, present in the words slogans drop+, support+ conclu+ initi+ [*desist+, apoi+ conclu+ inici+*]. In the second group (Classes 1 and 4), in turn, it became clear the meaning of the importance of monitoring the performance of the student by the mediator throughout the course and the strategies used in this, considered through the

analysis of the words slogans curs+, get+, not+ and long+ [*curs+, cheg+, not+* e *longo+*] in their elementary context units. It is clear that the elements of student motivation are their affective relationship with the mediator, assessment, and conclusion of the course.

> (…) my mediator was great. She was always attentive to everything and very supportive, even emotionally when difficulties arose. (*Ind_479 *Gen_F *Turma_DF02).

Monitoring and assessment of distance learning courses is a great challenge, especially when what is at stake is quality of education. The course adopted a process of training evaluation, therefore, it was necessary to set up a monitoring structure, in which the students were evaluated by the pedagogical mediator throughout their process learning, carried out a self-evaluation of their performance in the course and, finally, participated in institutional evaluations. According to Roque (2011), the evaluation and monitoring of distance-education students requires special considerations, since distance learning, when guided by pedagogical assumptions, aims at developing the critical autonomy of the students. Besides, distance education does not include the physical presence of the teacher, which makes necessary the development of work methods that promote student trust, enabling the elaboration of their own judgments and their ability to analyse them.

It is clear, from the statements of the students, that there is a connection between mediation and good performance in the course, even as a way to enable its conclusion. A relationship of motivation and affectivity through the encouragements and support in the realisation of activities also became evident. The item 'course conclusion' highlighted the importance of the pedagogical mediator in the completion of the course of each student. In their oral statements, the participants managed to show how the mediator motivated them in order for them to achieve their goals and reach the end of the specialisation.

> The pedagogical mediation carried out during the course contributed to my academic performance in all possible ways. It was extremely important for my following the course, with encouragement and support. (*Ind_780 *Gen_F *Turma_MG01).

The student's depictions revealed the importance of the cooperation dyad involving affectivity and monitoring.

The subdivisions of Class 2 were named Pedagogical Strategies (Class 1) and Knowledgebase Mediator (Class 4). They do not represent in themselves the explicit contradiction between their hierarchically superior class. The Class 2 set, in contradiction with Class 3, explains the importance of the mediator-learner dyad.

5 CONCLUSIONS

The testimonials of the students, who are teachers in the public school system, reveal the unfoldings of reflection in their professional practice. Pedagogical mediation, by encouraging the students to reflect, also promotes the transformation of their classroom performance, their education, and professional development. Pedagogical mediation was perceived by the students in the support they received during the realisation of activities and interaction in the proposed forums. Interestingly, the answers showed no reference to necessary physical presence of the mediator. Throughout the development of the course, pedagogical mediation appears as a permanent and individualised activity, whose objective is monitoring the students and offering each of them the help necessary in the learning process. Despite the fact that the central issue of this work has focused on evidences of the contribution of pedagogical mediation in the academic performance of students, looking into themselves, the actions were viewed as important to the improvement of professional performance – the main goal of the course – besides being highlighted as the reason for students to conclude the course.

Resuming the pragmatist perspective pointed out by Reinert, the object of analysis cannot be separated from its context. We emphasize, then, that the situation of the pedagogical mediator and the student is not static. The relationship of cooperation between them is dynamic, since it is alive and maintains direct and individual communication with each student. The three definitions of the word 'meaning,' namely, sensation or intuition, direction or motion, and significance, were demonstrated in this paper based on the mediator-learner dyad in the context of cooperation throughout the course, towards professional qualification and, finally, in the meaning of change that can be verified in the comments from the students.

ACKNOWLEDGEMENTS

The authors would like to thank Cilio Ziviani for his collaboration in sharing his expertise on Alceste.

REFERENCES

Benzécri, J. P. (1992). Correspondence analysis handbook. New York: Marcel Dekker.

Camargo, B. V. (2005). ALCESTE: Um Programa Informático de Análise Quantitativa de Dados Textuais (pp. 511-539). In A. S. P. Moreira, B. V. Camargo, J. C. Jesuino, & S. M. Nóbrega (Orgs.), Perspectivas teórico-metodológicas em representações sociais. João Pessoa, PB: Editora Universitária UFPB. Retrieved Apr. 23, 2013 http://www.laccos.org/pdf/Camargo2005_alc.pdf.

Deutsch, M. (1973). The Resolution of Conflict. New Haven, CT: *Yale University Press*.

Hegel, G. W. F. Science de La Logique - Premier Tome. Premier livre. L'Être. Edition de 1812. Paris: Aubier Montaigne. Traduction, présentation par P.J. Labarrière et Gwendoline Jarczyk.

Kalampalikis, N., Bauer, M. & Apostolidis, T. (2013). Science, technology and society: *the social representations approach. International Review of Social Psychology*, 26(3), 5-9.

Johnson, Burke. Quantitative, Qualitative, and Mixed Approaches Fourth Edition Educational Research, University of South Alabama. *SAGE Publications*, Inc. 2012.

Lima, Laura. Programa Alceste, Primeira Lição: A Perspectiva Pragmatista e o Método Estatístico. *Revista de Educação Pública*, Cuiabá, v. 17, n. 33, p. 83-97, Jan.-Apr. 2008.

Marchand, P. (2007). Concepts, méthodes, outils. Em C. Gauzente e D. Peyrat-Guillard (Orgs.), Analyse statistique de données textuelle en sciences de gestion: concepts, méthodes et applications (pp. 47-70). Paris: Édtions EMS.

Reinert, M. (2000). La tresse du sens et la méthode "Alceste" : application aux "Rêveries du promeneur solitaire". JADT 2000: *5es Journées Internationales d'Analyse Statistique des Données TeArtigo* - Referênciasxtuelles. Lausanne, Suíça.

Reinert, M. (2001). Alceste, une méthode statistique et sémiotique d'analyse de discours. Application aux "Rêveries du promeneur solitaire". *Revue Française de Psychiatrie et de Psychologie Médicale*, 49, p. 32-36. Retrieved Apr. 25, 2013 at http://www.uottawa.ca/academic/arts/astrolabe/articles/art0049/Alceste.htm.

Reinert, M. (2007). Contenu des discours et approche statistique. Em C. Gauzente e D. Peyrat-Guillard (Orgs.), *Analyse statistique de données textuelle en sciences de gestion: concepts, méthodes et applications* (pp. 21-45). Paris: Édtions EMS.

Roque, Gianna. Avaliação de aprendizagem em atividade desenvolvida a distância. In Campos, Gilda H.B. et al. As relações colaborativas: desafios da docência online. Curitiba, PR: CRV, 2011.

Using Activity Theory in Developing Instructial Tools for Project Management Studies

Maritta Pirhonen
Department of Computer Science and Information Systems, University of Jyväskylä, Jyväskylä, Finland
maritta.pirhonen@jyu.fi

Keywords: Project-based Learning, Project Management Education, Activity Theory, Supervising, Instruction.

Abstract: Competence and skills of the project manager are significant to project success. The skills needed in project managers' work cannot be learned only by reading the books or a lecture hall; one learns them by practice. Therefore, an important challenge for educational institutions is to develop pedagogical practices that allow students to participate in working life projects and to confront real-life problems. Project-based learning (PBL) offers a model that enables students to practice the skills and competences needed in working life projects by utilizing real-world work assignments in time-limited projects. Using PBL method alone does not necessarily guarantee learning result. In order to be successful, PBL method requires effective and competent supervision and guidance of students as well as appropriate tools for instruction. In this study the concepts from activity theory (AT) are applied to development tools for supervising project-based learning.

1 INTRODUCTION

The connection between project manager's skills and project success has been addressed in several studies (Iacovou and Dexter, 2004); (Müller and Turner, 2007). Often these skills are learnt in real-life working situations, because acquiring skills necessitates experience instead of studying theoretical facts by reading a book, or attending a lecture. Learning necessary soft skills required in IS project management and leadership during the project studies might support IS projects succeeding in "real world" working scenarios. Therefore, project management education needs to focus on practical issues of managing rather than on tools and techniques of management itself.

Project-based learning (PBL) offers a model for students to practice the skills and competences needed in IS projects by utilizing real world work assignments in time-limited projects (Tynjälä et al., 2009). However, using PBL method alone does not guarantee learning results. In order to be successful, PBL method requires both effective and competent supervision and an uniform learning environment that enables easy access and use of online materials.

The goal of the research in progress is to develop pedagogic methods and tools to support the learning of skills and competencies required in IS project work. Activity Theory (AT) concepts provide an analytical framework for developing the instructional methods responding to the educational needs. Particularly the concept of contradiction ("historically accumulating structural tensions within and between activity systems" (Engeström, 2001, p. 137) is seen to provide rich and fruitful insights into the system dynamics.

This paper is organized as follows. First, we depict the course, which is based on the project-based learning approach. Second, pedagogical background for project-based learning is reviewed. In the following chapter brief description of activity theory is presented. This is followed by the description of the project management course as an activity in AT. Finally, an outline of our ongoing research is presented.

2 PROJECT MANAGEMENT COURSE

The project-based learning (PBL) approach has been adopted in information system education at the University of Jyväskylä for years (see more Pirhonen 2009, 2010). For example, the implementation of the Project Management and Execution (PME) course (10 ECTS credits) is based

on PBL approaches (Tynjälä et al., 2009). The course belongs to the elective studies towards the degree of Master of Economic Sciences in the field of ICT. The main aim of the PME course is to offer the students an opportunity to gain authentic practical experience of an ICT experts' work. In addition, the goal is to provide students with a comprehensive and a realistic view of the work in IS projects. In more detail, students are expected to learn project management, leadership, group work, and communication by managing, leading and executing information systems projects. In addition, they are expected to learn an assessment of the significance of team leading as a part of project success.

The learning environment is maintained in co-ordination with three parties – a student group, the university, and a client organization. A legally binding cooperation contract is drawn up between the three parties before project starts. It covers the subject matter (a description of the project objectives), the obligations and rights of the contracting parties, copyrights, guarantees and maintenance, confidentiality and the concealment of confidential information, payments and the payment schedule.

The project course lasts from the beginning of November to the end of April (26 weeks). During the course each student is expected to use 140 hours for implementing the project task and 130 hours for demonstrating project-work skills, including team leading, group work, and communication. The groups plan their work, complete the scheduled tasks, and produce deliverables. Each student is expected to take the role of project manager and project secretary. These roles rotate every month to ensure that each member of the project team works in both roles at least once. In total, a group of five students uses 700 hours in planning and executing the client project.

During the course, students work in close co-operation with their client and they meet with the client representatives on weekly basis. In addition, the guest lectures from collaborative companies are invited to give lectures on relevant topics to project management. The collaboration with a client ends in a steering group meeting at which the results of the project are approved.

During the course seminars are arranged to enhance students´ communications skills.. Pedagogical activities such as peer reviewing, group discussion, peer coaching and self/peer assessment are being set up by supervisors to enhance the learning effectiveness of the project.

Each student group is evaluated twice during the six-month period of the project. The first evaluation takes place in the middle of February after three months' work. The second evaluation is carried out at the end of the course in April. The content of the evaluation is grouped and structured around the themes covering issues to the course's learning objectives and critical to project management success. The course grade (1-5) for a group is calculated on the basis of the following factors: project management, project work, and communication. The evaluation involves composing an evaluation report using the assessment framework. Both students and their supervisors compose the report. Written evaluations are uploaded in the digital learning environment Optima (the day before an evaluation discussion). The evaluation is based on the perceptions of the students´ work capabilities with their clients as well as the documentation produced during the project. Both supervisors and the student groups are acquainted with the each other's evaluations before an evaluation discussion. The grading of the course is mainly based on the debates that emerge during discussions concerning the reports.

3 PROJECT-BASED LEARNING

Problem-Based Learning (PBL – nowadays also the abbreviation for Project-Based Learning) has become widely recognized over the last 40 years. PBL has been proven to be successful educational approach in many different study domains. It has been adopted for years in Aalborg University in Denmark (Graaff and Kolmos, 2003). According to Kjersdam (1994) students graduated from Aalborg are more productive and competent compared to graduated students from other educational institutions.

Project-based learning (PBL) refers to a theory and practice of utilizing real-world work assignments on time-limited projects to achieve mandated performance objectives and to facilitate individual and collective learning (Smith and Dods, 1997). The theory of PBL is based on constructivism and according to the constructivism theory, the learner is guided to build and modify his or her existing mental model. This means that the focus is on knowledge construction rather than on knowledge transmission as in the theory of behaviourism. Constructivism takes account of the situational nature of learning and thus advocates authentic or simulated environments (von

Glasersfeld 1984). There are five significant features of PBL (Helle, Tynjälä, Lonka and Olkinuora 2007):

- a problem or a question serves to drive learn-ing objectives;
- a concrete artifact is constructed;
- the learners control the learning process (pacing, sequencing, and actual content);
- the learning is contextualized (what we learn in a particular context we recall in similar contexts); and
- projects are complex enough to induce students to generate questions of their own.

In many models of project-based learning, students are assumed to work on real world projects by default. This creates good conditions to learn a vast range of skills in various project areas. Students learn management, teamwork, and communications, as it involves both individual and co-operative activities, interactive discussions and writing in the form of plans, reports, memos etc. This type of learning offers a very concrete and holistic experience of certain processes such as the process of construction work or managing a project (Helle et al., 2006). Often collaboration skills are put into action by the collaborative nature of project management. In fact, the studies have suggested that project work may have many educational and social benefits (Moses et al., 2000), such as the development of communication skills (Pigford, 1992), along with team-building and inter-personal skills (Roberts, 2000). Supervisors support the work of their students by guiding and assisting them to learn independently and helping them to retrieve relevant information when required. Supervisors oversee the project process and monitor the progress and performance of each student. The role of the supervisor is vital, especially in the early stages of the project when students may need more guidance in situations where they need to communicate and collaborate with their client.

4 ACTIVITY THEORY – AN OVERVIEW

Activity theory (AT) offers a theoretical framework to study both individual and collective activities. It provides an analytical framework within which to study human activity in general. A model of the structure of an activity system (AS) includes two types of constituents: core components, such as subject, object/outcome, and community; and mediatory components, such as instruments (tools), rules, and division of labor (Figure 1).

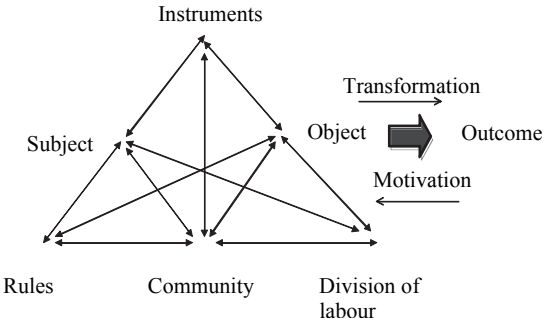


Figure 1: The structure of a human activity (adapted from Engeström, 1987, p. 78).

An activity is a collective phenomenon; it has a subject (an individual or collective) who understand its motive, and who uses tools to achieve an object, thus transforming objects into outcomes. An activity is always associated with long-term purposes and strong motives. All members of the community share the object (and the motive) of the activity. Tools mediate between a subject and the object, which is transformed into the outcome. The object is seen and manipulated within the limitations set by the tools. Rules mediate the relationship between the community and the subject, while the division of labor mediates the relationship between the community and the object. Rules cover both implicit and explicit norms, conventions, and social relations in a community as related to the transformation process of the object into an outcome. The responsibilities of the members of the community are coordinated by some division of labor (e.g., the division of tasks and roles among members of the community and the divisions of power and status), yet guided by rules. These rules regulate, as well as constrain, their actions and relationships in the activity system (Engeström, 1990; Kuutti, 1996).

Engeström (1987) added the concept of contradiction onto Vygostky´s (1978) thinking. Primary contradictions are those found within a constituent of the activity (i.e., in the object, rules, tools, etc.) and secondary contradictions are those that appear between constituents of the activity (e.g., between the tool and the subject). Contradictions constitute a key principle in AT and shape an activity (Engeström, 2001). When contradictions arise, or when they are observed, they expose dynamics, inefficiencies, and importantly, opportunities for a change (Helle, 2000). Kuutti (1996, p. 34) describes contradictions as “a misfit within elements, between them, between different

activities, or between different development phases of a single activity". They generate "disturbances and conflicts, but also innovative attempts to change the activity" (Engeström, 2001, p. 134).

Contradictions are significant for development and they exist in the form of resistance to achieving goals of the intended activity. They also exist as emerging dilemmas, disturbances, and discoordinations. In spite of the potential of contradictions to result in development in an activity system, the development does not always occur. Often contradictions may not be easily recognized or acknowledged, visible, or even openly discussed by those experiencing them (Engeström, 2001). On the other hand, contradictions that are not discussed may be embarrassing, or uncomfortable in nature. They may also be culturally difficult to confront, such as personal habits, bad behaviour, or an incompetence of the leader.

To summarize, subjects, who are motivated by an object, carry out activities. A subject transforms the object into an outcome. An object may be shared by a community of people, working together to achieve a desired outcome. Tools, rules, and a division of labor mediate the relationship between the subjects, community, and the object. Contradictions are a key principle in AT and they are driving force of change.

5 PROJECT MANAGEMENT COURSE AS AN ACTIVITY

In the depiction of PME course as an activity system, a student group is chosen to be subject. A subject plays a key role when analysing other elements of an activity. In this case we are interested in student group´s perspective – how the tools support their learning and achieving the course's learning objectives - when analysing PME activity. The objectives of the cooperation parties differ. From the students' and supervisors' points of view, the main object is to learn useful skills needed in "real project work". Correspondingly clients' main motive to co-operate with the university and by doing so, find potential employees to recruit. Certainly, clients' objective is also to obtain results from the project they are involved with. This is a goal they naturally share with the students they have worked with. Different types of objectives, however, might cause contradictions between the parties involved. If such an event occurs, the supervisors need to intervene in the situation by discussing openly about the issue with all parties. In our study we focus on the objective seen from the point of student group´s view. Their motive is to achieve the course's learning objectives. The outcome of course of is that students are provided with skills needed in projects. The activity "PME course" is presented in the terms of activity theory (AT) in Table 1.

Table 1: PME course as an activity.

5	Description
Subject	Student group
Object	To learn skills needed in project management
Outcome	To enable students to develop skills needed in project
Instruments or tools	Project management tools, communication tools, guiding meetings, written instructions, pedagogic methods
Community	Students, teachers, clients
Division of labour	Responsibilities according to the contract
Rules	Constraints on schedule, contract, assessment

The tools include a project management system (e.g. software, standards), weekly meetings between the supervisor and the student group, and written instruction. During the meetings, the weekly project reports and project plans are discussed and reviewed. The project manager and team members keep providing updates of their project, which are compared with the documented expectations in the project plan.

The present tools are seen to have troublesome features. First, the amount of the tools required in the project work is great, and they are located in several, different environments. Project management tools (e.g. software for managing the schedule or the resources) are located in multiple systems and data transmission between systems has been proven to be difficult and time-consuming. Second, the written instructions are stored in the digital learning environment into which students need to log in separately in order to gain access to project documentation, or upload and share new documents. Students are frustrated while working with so many incompatible systems, which may even decrease their motivation to study.

Activity theory emphasizes that a tool should come fully into being when it is used and that knowing how to use it is a crucial part of the tool. Therefore, the use of tools entails an evolutionary accumulation and transmission of social knowledge, which influences not only the external behaviour but also the mental functioning of individuals. Therefore, the pedagogic tools supporting and promoting learning are vital part of supervisors work with their students.

6 CONCLUSIONS

Effective and competent supervision and guidance of students is a vital part of a project-based learning method; PBL method alone does not guarantee learning results. Hence, appropriate pedagogic instructional tools and methods are of critical importance of achieving learning goals.

To understand the underlying contradictions between a student group and tools used in project studies, we adopt the activity theory (AT) as our lens to explore possible misfits. The strength of AT is that it allows to break down the structure of an activity into smaller categorical elements (Basharina, 2007), and to identify contradictions and structural tensions of the activity (Engeström, 1995); (Engeström, 2001). Contradictions relate to tendencies or forces that need each other, but at the same time negate each other. The contradictions generate disturbances, conflicts, and eruptions in an activity, thus making contradictions indirectly visible. By recognising structural tensions that causes disturbances and conflicts in activity it is possible that new forms and qualitative stages of activity emerge as solutions to the contradictions (Engeström, 1987). This being the case, we argue that the AT provides us with the proper theoretical lens to develop instructional tools for project management studies at the University of Jyväskylä.

So far we have modelled the PME course as an activity system. Next step in our study is to start an exploratory study by interviewing students, supervisors, and clients having participated in the PME course in 2011 - 2014. The aim of the study in progress is to identify the disturbances emerged during the course and contradictions that cause "problems, ruptures, breakdowns, and clashes" (Kuutti, 1996, p. 34). In this phase of the study we are especially focusing on contradictions found between the student group (subject) and pedagogic methods and tools used during the course. Further studies may also benefit from a deeper investigation of the objectives of the PME course from clients' points of view for purposes to find contradictions between different objectives of the cooperation parties.

ACKNOWLEDGEMENTS

I wish to thank Eliisa Jauhiainen and Minna Silvennoinen their insightful feedback in the development of this study.

REFERENCES

Basharina, O. K. 2007. An activity theory perspective on student-reported contradictions in international tele-collaboration. *Language Learning & Technology,* 11(2), 82-103.

Cole, M. and Engeström, Y. 1993. A cultural-historical approach to distributed cognition. In G. Salomon (Ed.) *Distributed cognitions: Psychological and educational considerations.* Cambridge University Press, New York.

Engeström, Y. 1987. *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-konsultit.

Engeström, Y. 1990. *Learning, working, and imagining: Twelve studies in activity theory.* Helsinki: Orienta-konsultit.

Engeström, Y., 1995. Objects, contradictions and collaboration in medical cognition: an activity-theoretical perspective. *Artificial Intelligence in Medicine*, 7 (5), 395-412.

Engeström, Y., 2001. Expansive learning at work: towards an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133-156.

de Graaff E. and Kolmos, A., 2003. Characteristics of problem-based learning, *International Journal of Engineering Education*, 19(5), 657–662.

Helle, M. 2000. Disturbances and Contradictions as Tools for Understanding Work in the Newsroom, *Scandinavian Journal of Information Systems,* 12(1), 81-113.

Helle, L., Tynjälä, P. and Olkinuora, E. 2006. Project-based learning in post-secondary education – theory, practice and rubber sling shots. *Higher Education*, 51(2), 287–314.

Helle, L., Tynjälä, P., Olkinuora, E. and Lonka K. 2007. 'Ain't nothing like a real thing'. Motivation and study processes on work-based project course in information systems design. *British Journal of Educational Psychology*, 77(2), 397-411.

Iacovou, C. L. and Dexter, A. S. 2004. Turning around runaway information technology projects. *California Management Review*, 46(4), 68-88.

Kjersdam, F. 1994. Tomorrow´s Engineering Education – The Aalborg Experiment. *Journal of Engineering Education*, 19(2), 197-204.

Kuutti, K. 1996. Activity Theory as a potential framework for human-computer interaction research. In B. Nardi (ed.) *Context and Consciousness: Activity Theory and Human Computer Interaction*, MIT Press, Cambridge, 17-44.

Moses, L., Fincher, S., Caristi J., 2000. Teams work (panel session) in Haller S. (ed.) *Proceedings of the thirty-first SIGCSE technical symposium on Computer science education*, March 7-12. Austin, USA. New York: ACM Press, pp. 421-422.

Müller, R. and Turner J. R. 2007. Matching the project manager's leadership style to project type. *International Journal of Project Management*, 25(1), 21-32.

Pigford, D. V., 1992. The Documentation and Evaluation

of Team-Oriented Database Projects. *Proceedings of the twenty-third technical symposium on Computer science education*, Kansas City, Missouri, United States. New York: ACM Press, pp. 28-33.

Pirhonen, M. 2009. Challenges of Supervising Student Projects in Collaboration with Authentic Clients. *Proceedings of the 1st International Conference on Computer Supported Education (CSEDU)* [CD-ROM], Lisbon, Portugal, May 23-26, 2009.

Pirhonen, M. 2010. Learning Soft Skills in Project Management Course: Students´ Perceptions. In Aramo-Immonen, H., Naaranoja, M. and Toikka, T. (eds.) *Proceedings of Project Knowledge Sharing Arena, Scientific Track Project Days* 2010, 31-41.

Roberts, E., 2000. Computing education and the information technology workforce. *SIGCSE Bulletin* (32)2, pp. 83-90.

Smith, B., Dodds, R., 1997. *Developing Managers Through Project-based Learning*. Aldershot/Vermont: Gover.

Tynjälä, P., Pirhonen, M., Vartiainen, T. and Helle, L. 2009. Educating IT project managers through project-based learning: A working-life perspective. *The Communications of the Association for Information Systems,* Vol. 24, Article 16, 270-299.

von Glasersfeld, E. 1984. An introduction to radical constructivism. In Watzlawick, P. (eds.) *The Invented Reality. How do We Know What We Believe We Know? Contributions to Constructivism*, Norton, New York, NY, 17-40.

Vygotsky, L. S. 1978. *Mind in Society. The development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.

Transforming Experience of Computer Science Software Development Through Developing a Usable Multiplayer Online Game in One Semester

Ilmi Yoon[1] and Eun-young Elaine Kang[2]
[1]*Computer Science Dept, San Francisco State Univ. 1600 Holloway Ave, SF, U.S.A.*
[2]*Department of Computer Science, California State University, Los Angeles, U.S.A.*
ilmi@sfsu.edu, eykang@exchange.calstatela.edu

Keywords: Multiplayer Online Game, Game Development, Industry Style Team Work, Project-based Learning.

Abstract: We present an instructional design of computer science project-based course to transform students' experiences of acquiring software development skills. In a collaborative classroom emulating a typical industry work setting, students will collectively create and build a Multiplayer Online Game using a variety of complex software components. A course was taught to design and develop a working Multiplayer Online Game within one semester: building a ready-to-usable game in one semester with whole classmates presents significant challenges to cope with and stimulate students to realize the important aspects of teamwork and software engineering principles. Students present their progress, discuss future milestones and trouble shoots, update documents for clearer communication and utilize source control tool throughout the semester. Unlike usual class setting, all students worked collaboratively together like one company to achieve the goal. In the class, students started from concept design and developed specific components of working Multiplayer Online Game, while broadly learning game design, 3D graphics, Game Engine, Server-client architecture, Game Protocol, network programming, database, Software Engineering principles, and large application development as a team project. The course was successfully transferred to CSULA in Fall Quarter, 2013.

1 INTRODUCTION

Over the past decade, the software development process has become exceedingly complex (Fuggetta, 2000). The industry now expects software engineers and computer scientists to excel not only in the technical aspects of software development, but also in self-learning of advances in new technology, the fluent communication of their ideas in presentations and documentation, often produced in a collaborative environment (Dutoit and Bruegge, 1998). Because software development today cannot be accomplished without collaboration across teams, experience in this area is highly desirable in computer science students entering the job market (Mencher, 2003; Parberry et al., 2005). Given the rapidly changing nature of technology, software developers can no longer rely exclusively on concrete knowledge previously attained. Rather, they must have the ability and adaptability to adjust swiftly to changing practices and emerging platforms. These abilities are seldom addressed in an academic setting, leaving new graduates minimally prepared and struggling in these important areas. There is an enormous need for courses that provide students with the training and experience to meet the changing expectations of industry and to successfully transform knowledge they have acquired in the classroom into the ability to produce complex, innovative, high-quality software products in a collaborative environment.

The pedagogy currently used in most Computer Science classes with team projects employs team projects of a few weeks to a whole semester in duration with several team members, involving minimal or no collaboration across teams. While such an approach is essential and effective for teaching specific content knowledge and achieving particular learning outcomes, it is also limited in that it focuses on the development of relatively simple programs to teach specific content knowledge and on the completion of classroom-specific assignments

that are not readily transferrable to an industry setting and do not adequately prepare students to work on large-scale programming projects in a highly synchronized work environment that demands complex contributions across teams. Students are rarely exposed to inter-team interactions or to working collaboratively across teams to develop a common product. Ultimately, the scope of the software developed in the typical CS classroom and the integration of its components tend to be simplistic as a result of the aforementioned limitations (Parberry et al., 2005). This paper presents an alternative approach to prepare students for the challenges they will face in industry: a course designed around a complex project to be completed by the class as a whole, giving students practical experience with inter-team collaboration and an understanding of the complexities of the software development process.

Objectives of the course are to: (1) teach students effective communication, presentation and collaboration skills that are often neglected in the traditional classroom; (2) motivate students to learn essential CS core content through peer interaction and encouragement; (3) enable students to produce a robust end product; and (4) give students practical insights into the real-world work environment to prepare them to master the challenges of a career in CS.

In next section, we discuss backgrounds on selecting Multiplayer game as instructional material. In section 3, we discuss the course design and structure to teach students to build a working MMORPG game within a semester (15 weeks). In section 4, we present the results and learning outcome. Conclusion and on-going work are discussed at section 5.

2 BACKGROUND - INSTRUCTIONAL MATERIAL

Massively Multiplayer Online Role Playing Game (MMORPG) is a type of online multiplayer game that allows thousands of players to play concurrently in a virtual world. Even after players log out or disconnect, the virtual world persists and the updates made by player remain. Such games form an online community of players who interact with one another and progress through the game over time. The quality of social interactions and character progression within the game radically increase the excitement of the game; popular games like WoW (World of Warcraft) and Second Life easily develop communities of multimillions of players, to the point that they have become significant trends in global youth culture (Song, S., Lee, J., 2007). It is the author's experience that being able to understand the behind-the-scenes work that goes into the creation of such games and then developing a game themselves strongly motivates students to focus on and excel in their CS coursework.

Social gaming and MMORPG are relatively recent game genres as they bloom over the availability of Internet. Unlike stand-alone game, a MMORPG game keeps virtual world persistent even after players log out, so game server has to run infinitely connecting all the clients and updating game status at database. Therefore, building a MMORPG game covers broad spectrum of computer science technology from computer graphics, 3D modeling, game engine, network programming, client-server architecture, and database, so it helps students exposed to diverse technical components comprehensively. Also a Multiplayer Online game is indeed a large application with a lot more challenges in debugging and quality assurance. Succeeding in building a reliable one ensures students to learn the qualities that software companies want.

3 IMPLEMENTATION

This class was designed for Computer Science senior and graduate students who completed programming languages, data structures and Software Development Principles but not necessarily have taken elective courses such Software Engineering, Network nor Database courses. Considering that majority of students find jobs at industry after graduation, the class was organized in more of industrial flavor; a team of students (1) receive tasks and milestones, (2) achieve the milestones by actively looking for solutions from any available resources, (3) interact with other teams collaboratively, (4) and produce documents for clear communications between teams and future extensions.

The class was organized into Game Concept Design team, Game Server team, Game Client team, Game protocol team, Database team, Art Support team, Game Content team, Testing team, and Launching team. Within first 2 weeks, objectives and the responsibility of each team were discussed and then students were assigned into one or two teams based on their interest and backgrounds. From

then on, teams worked on in parallel to achieve their milestones, while learning all the required technology for the given task and sharing their learning with other teams.

Sample Summary of team tasks is as below.

Game Concept Design Team: this team has most important initial responsibility. The objective of the whole class was to build an educational MMORPG for Computer Science students, especially for beginners who are taking first programming courses. The concept design team should consider the limitations such as time (10 weeks for actual implementation), lack of artists, and unknown factors like the potential of other team's performance. The team also has to collect everyone's idea, finalize the game concept into several milestones and produce the documentation.

Game Client Team: this team is responsible to implement 3D game client using Panda3D. Students have to learn 3D graphics, Panda3D game engine and Python. Also client communicates with server using the game protocol to update the status (position, motions, and levels, etc) of other clients concurrently in the same virtual space.

Game Server/Protocol Team: this team is responsible to implement game server that connects to all the clients, synchronize every client status and update the Database. Team did not start from scratch as there was a JAVA game server written for a similar MMORPG game called "NurseTown" by students of Internet Application course. The server team had to read and revise the game server to work for "deBugger" game. Also the server team has to develop an effective set of protocols used between clients and server. Students in this team have to learn socket programming, client-server architecture, and Network protocols.

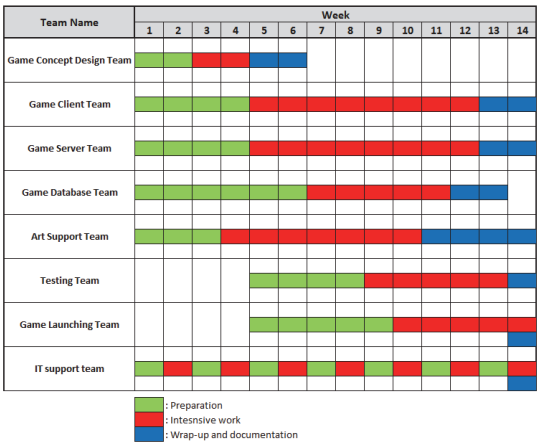


Figure 1: Timelines for each team over weeks during the semester.

Art Support Team: this team is responsible to provide 3D models to create the virtual world and characters in the game. As there are no artists in the class, this team has to find freely available 3D models and convert the format to work with Panda3D. Students in this team have to learn basic of 3D modeling tools such as Blender or Maya and converting 3D model formats, and real time rendering performance trade-off (visual quality vs. rendering speed).

Game Content Team: "debugger" game is a unique MMORPG game for educational purpose. Therefore, the educational game contents are as important as other game assets. This team develops educational contents. As senior or graduate students of computer science, this team is already familiar with educational contents that beginner programmers learn which is good and bad. This team needs to interview beginner students, observe the troubles that they probably forgot and produce effective educational contents.

There were a little over 20 students in the class, so each student choose one primary team and one secondary team, so each team has at least 2 to 3 students. By having students in a few different teams, students could learn more than one technical area and teams can communicate and update the progress easily. Each team has to work very collaboratively as the whole class works for one project together.

Important concepts and technologies were taught to the whole class by several lectures, but most of class meetings were used as presentation of milestones, progress, trouble shootings of problems on hand, so the whole class stays in a same page and gets exposed to the problem and solutions.

Software Engineering principles have been practiced throughout the whole course as teams realized the importance of the inter-team communications due to the rapid developments in parallel. Changes in one team did propagate changes in milestones or requirements in other teams. SVN was actively used.

4 RESULTS

The whole class successfully developed a working MMORPG game at both 2009 Fall (DeBugger - http://smurf.sfsu.edu/~debugger/), 2011 Fall (World of Balance - http://smurf.sfsu.edu/~wob/) when the courses were offered in this style, 2013 Spring (extension of World of Balance game), and 2013 Fall (extension of DeBugger game, California State Univ. Los Angeles). The games produced by the

class had intention to be used for educational game research or science discovery game research which added substantial challenges (entertaining while achieving another purpose) compared to pure entertainment games. As a whole class, students have constantly struggled together to validate the game design for achieving its intended purposes. Developed games are currently being used as intended (Yoon et al, 2011; Yoon et al 2013)

DEBUGGER GAME (Educational MMORPG For Computer Science Students)

Each team achieved milestones close to the initial (very ambitious) goal. **Game concept team** came up with the title, "deBugger" and the theme of the game; fighting against bugs within inside of computer, inspired by the origin of the word and the significance of the word to computer science students. Team also prioritized concepts in to several milestones, so kept the milestones doable for one semester, but left many interesting ideas for future extensions.

Primary milestone for Fall 2009 was to build virtual world where players can explore and fight with bugs by solving multiple choice questions. Health drops under bug's attack and game items can be used to shield from the bug's attack (delay the health damage or clear out options from the multiple selections). Players can level up by solving questions required by each level. Also player should be able to chat, create friends' list and maintain their inventory (trading or giving gifts).

Art Support Team achieved beyond the requirement. Students learned 3D modeling tools and created 3D environments for the game. We used 3D character models from Panda3D repository, but all the environments (figure 2) were created by art support team and theme was inspired by movie Matrix and Tron during inspirational discussions at class.

Client Team was the largest team and they were divided into sub teams to be responsible for specific tasks in parallel. List of such tasks are Registration Process (Avatar Selection), Login Process, Chat, Bubbles (translucent panel), Chat Bubbles, Character Name and Damage Bubbles, Friends, Inventory, Character Info, Hot keys, Camera control, Character Movements, Interface (mouse, Keyboard), Collision, Battle System, NPCs (Non Player Character), Bug (with AI), Mini maps Client team also achieved impressive progress of completing all the requirements and the resulting game client was reliable and capable of very smooth rendering. Each task listed above took intensive effort to understand the technology (3D graphics) to manipulate Panda3D game engine.

Server team worked hard to make MMORPG game alive. Server team was able to reuse the majority of server code built for a game called, "Nursetown" by Prof. Ilmi Yoon and her graduate students. Server team had to understand the communication mechanism and modified DB schema and added lots of new game protocols.

Game Contents team focused on adding educational contents into the game. Game contents team structured pedagogy of computer programming learning and created quizzes in different levels for such purpose.

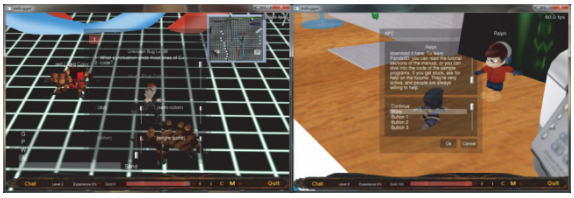


Figure 2: Left: A Scene from the game: NPC (Non Player Character) requires simple AI to interact with players; guide players to follow rules of game or inform about events to participate. Right: A scene from the game: Bugs are attacking players. Bugs have AI to chase the player, handle collision and re-spawn.

Data Base team studied design of efficient DB schema and developed it for debugger game. Team also installed mySQL server and deploy the data sets.

Testing team was developing a test client that has a minimal AI to wander around, acts like a usual client and save the communication log, so we can use the test client to measure the scalability of the server. In the middle of the semester, we found that the debugger game needs a bug server that controls all the bugs that has AI to attack the player and wander around the debugger world. Testing team took the role and successfully completed the creation of the bug server with the close collaboration with the client team.

Launching team studied popular MMORPG hosting company's web site and designed a web site for the possible full version of the debugger web site and implemented core set of it to be able to launch the game for programming class students (figure 13). The launching team made a simple web site (http://smurf.sfsu.edu/~debugger) with link to download client application (self-extracting installer for windows), game rules, information of how to get started, and screen captures.

The class as one whole team completed a working game (version 0.9) by the end of semester. The game contained core MMORPG game features.

The client and server connect and communicate in stable manner. DB stores persistent data. User can create an account with options of a few characters. And then user can log in at the location where they logged out last time. Player can explore the virtual world starting from a desktop space into the inside computer world. User can find if friends are logged in and chat with friends or other players within the same space. Players can check their game items at the inventory box, health bar and the current level as well as their current position at the mini-map. Most importantly, players can battle with bugs using the multiple choice questions and the health bar and level were properly updated.

At the same time, whole class produced one giant document (437 pages, available at http://smurf.sfsu.edu/~debugger/download_documentation.php) of each team's work for future developers who will extend this game to the next level.

At the Fall 2013, this course was adapted at CSULA for the first time and the students extended the DeBugger game by adding multiple mini games (http://csproject.calstatela.edu:8080/debugger/#/). Due to the transition from the semester system to quarter system, there were additional challenges. Even so, the students' evaluation shows the effective learning in large scale projects and communication skills.

WORLD OF BALANC (Science Discovery Game for Computational Ecology Research)

This game focuses and extends the scope of game from simple entertainment or education further into Gamification or crowd computing that enables scientific data collection through game play. While players are playing an ecosystem nurturing game, they are also producing data that can be meaningfully used by ecologists (Yoon 2013). Details of game can be found at the given web site (http://smurf.sfsu.edu/~wob/). This game project presented additional challenge of understanding computational ecologists' research agenda to be able to develop effective science discovery game. The content team also worked extra hard to connect the game with computational engine that scientists use for their research. Over two semesters, game engine (client side) has been changed from Panda3D to Unity3D to support multiple platform (iPad, Windows, etc). The game has been used for learning/training impact study and play test with 20 players for a week.



Figure 3: A Scene from World of Balance Game (Unity3D Client), showing the biomass of the species within the game.

5 CONCLUSION AND ON-GOING WORK

Students were proud of what they achieved in one semester (15 weeks). Considering that the first few weeks were used to the introduction to the class and setting up the team, students had about 2.5 months to work on the game. To be able to make the MMORPG game work, each team has to deliver what they were expected in milestones. It serves as a good pressure on each team and students collaborated across team enthusiastically. Most students put efforts beyond requirements, resulting in good learning outcome.

This class can serve as a good team project course for senior students in Computer Science, so they can taste the industry style problem solving approach and team work in academic environment. Also as a by-product, the class produced a usable game that is currently used in CSc 210 Introduction to Computer Programming course at SFSU. The objective of using this game for educating CS freshman seemed to give pressure to students to deliver a smoothly working game as their clients are sitting right next to them, so it lets them alert about their end user.

The project's focus is not only producing the end product, but assessing/evaluating students' learning from the experience. We have created rubrics to measure students learning (Table 1) and plan to assess and evaluate them thoroughly in Spring 2014.

To be able to replicate the same class in the future, course material needs to be better collected and organized. There are hundreds of discussion threads at iLearn news group that were effective while being used, but not so organized, so may not be transferrable to next year class nicely. Inter-team communications were well captured to be used in the class for each team progress presentation and documentations. But inner team communication and resources were not saved. Each team created very

useful resources to be shared within team to learn the required technology for each team for example, client team learned Panda3D, Python, collision handling together helping each other. Same applied for each team. Vast amount of self study materials were created by each team. These can be better managed to be created at wiki, so they can produce more re-usable study resources.

ACKNOWLEDGEMENTS

This project is being funded by National Science Foundation Div. Of Biological Infrastructure, Biological Databases and Information, NSF DBI- 0543614. Also this project is recently funded by NSF Transform Undergraduate Education (TUES) 1140939 for 3 years to prove the learning efficacy quantitatively and spread widely.

REFERENCES

Dutoit, A. H., Bruegge, B, 1998 "Communication metrics for software development," Software Engineering, IEEE Transactions on, Vol. 24 Issue: 8, pp. 615 – 628.

Fuggetta, 2000 "Software process: a roadmap," ICSE '00 Proceedings of the Conference on the Future of Software Engineering, ISBN:1-58113-253-0.

Jones, R. M., 2000 Design and implementation of computer games: A capstone course for undergraduate computer science education. In Proceedings of the 31st SIGCSE Technical Symposium on Computer Science Education, pages 260–264. ACM Press, 2000.

Mencher, M., 2003 "Get in the Game!," New Riders Publishing.

Moser, 1997 "A fantasy adventure game as a learning environment: Why learning to program is so difficult and what can be done about it." In Proceedings of the 2nd Conference on Integrating Technology into Computer Science Education, pages 114–116. ACM Press, 1997.

Parberry, 2001 "Introduction to Computer Game Programming with DirectX 8.0." Wordware Publishing, 2001.

Parberry, I., Roden, T., Kazemzadeh M.B., 2005 "Experience with an industry-driven capstone course on game programming: extended abstract," Proceedings of the 36th SIGCSE technical symposium on Computer Science Education, Vol. 37 Issue 1.

Song, S., Lee, J., 2007 "Key factors of heuristic evaluation for game design: Towards massively multi-player online role-playing game, International Journal of Human-Computer Studies, no. 65, 709-723.

Yoon et al. 2011 Ilmi Yoon, Gary Ng*, Zoran Millic*, Supakit Kiatrungrit*, Yiyi Miao*, and Sunggye Hong, "Educational Multiuser Online Game, 'DeDugger' Game for Introductory Computer Science Class," International Conference on Frontiers in Education: Computer Science and Computer Engineering, July, 2011, pg. 393-398.

Yoon et al. 2013 Yoon, I., Ng, G., Rodrigues, H., Nguyen, T., Paik, J., Yoon, S., Williams, R., Martinez, N., "Iterative Design and Development of the 'World of Balance' Game: From Ecosystem Education to Scientific Discovery," IEEE International Games Innovation Conference (IGIC), Sept. 2013, Vancouver, British Columbia, Canada, Pg. 283-290.

Yoon et al. 2013b Ilmi Yoon, Arno Puder, Gary Ng*, Manori Thakur, Hunvil Rodrigues, Jae H. Paik, Eun-Young Kang, "Educational MMORPG for Computer Science: DeBugger, a Virtual Lab on PC and Smart Phones," International Workshop on Distance Education Technologies, Brighton, United Kingdom, August, 2013, pg. 96-10.

APPENDIX

Table 1: Project goals and corresponding assessment mechanism.

Goal 1: Teach students effective communication, presentation and collaboration skills that are often neglected in the traditional classroom.	
Assessment Rubric	Assessment Mechanism
• Students presentation skills both oral and written • Ability to communicate complex CS concepts • Students ability to work in groups • Organizational skills • Focus of task • Students ability to communicate clearly via exchange of shared documents and protocols with team members	• Students evaluation of skills and competencies via survey • Interviews and focus group discussion • Classroom observations (via a protocol) • Records of public presentations/journal publications at conferences and professional meetings

Table 1: Project goals and corresponding assessment mechanism. (Cont.)

Goal 2: Motivate students to learn essential CS course content through peer interaction and encouragement, thereby increasing course engagement and making learning experience relevant and fun	
Assessment Rubric	Assessment Mechanism
• Student reflections of their roles • Student reflections of their own learning process • Students skills and strategies to engage in the design and development process to create and engaging product.	• Students' course satisfaction survey • Interview and focus group discussion • Student self-evaluation
Goal 3: Enable students to produce a robust end product, i.e., a game that can be played by the public and that can be continuously expanded by other students in successive classes	
Assessment Rubric	Assessment Mechanism
• Students design and development skills • Students critical thinking skills applied to the development of the game • Student content knowledge • Originality of ideas	• Performance evaluation of the game design process • Interview and focus group discussion • Protocol to measure effectiveness of the product (MMORPG game) developed • Student self-evaluation
Goal 4: Give students insights into the real-world industrial work environment to prepare them to master the challenges of a CS career	
Assessment Rubric	Assessment Mechanism
• Effectiveness of the class in providing hands-on experience and skills to succeed in CS career. • Students' determination to pursue a CS degree • Long-terms career goals in CS or other STEM field	• Interview and focus group discussion • Student on-line surveys

r(e)flect
The Reflective Teaching Material about Energy, Behaviour and Product Development

Magdalena Boork[1], Susanne Engström[2], Rebekah Olsen[3] and Therese Balksjö[3]
[1]*SP Technical Research Institute of Sweden, Energy Technology, JTI, Box 7033, SE-750 07, Uppsala, Sweden*
[2]*Department of Education, Uppsala University, Box 2136, SE-750 02, Uppsala, Sweden*
[3]*Interactive Institute Swedish ICT, Energy Design, Portgatan 3, SE-633 42, Eskilstuna, Sweden*
magdalena.boork@sp.se, susanne.engstrom@edu.uu.se, {rebekah, therese.balksjö}@tii.se

Keywords: Teaching Material, Reflection, Visualization Tool, Energy Use.

Abstract: r(e)flect is a tangible curriculum kit for students age 10-15 to reflect on energy behaviour and make better informed choices about energy use. Along with web-based material, the kit includes a minicomputer, smart plugs (sensors), and an electricity visualization tool especially designed to be used in the classroom to conduct experiments and measurements, perform project work in product development, and increase the understanding of the kWh concept. The curriculum is closely connected to the new Swedish National curriculum and supplies the teachers with appropriate support for assessment of different skills. The project was initiated in 2011 and a first version of the curriculum was tested and evaluated with students and teachers. The feedback from this trial influenced the second iteration of the kit. The new version will be tested in at least 20 schools during the spring of 2014 and this second phase will continue until 2015. The physical r(e)flect material can be borrowed by teachers free of charge, and the web-based platform is open and accessible to anyone.

1 INTRODUCTION

In Sweden, a new national curriculum including course plans in physics and technology was implemented in 2011. Additionally, a new grading system with clearer criteria was introduced. In physics and technology, the subject content and what should be assessed are clearly stated. Within physics, concepts such as energy and power should be dealt with, students should develop an understanding, but also be able to discuss, debate and reflect upon their own energy use. Within technology, students should get insight into the design and development process and their creative and entrepreneurial abilities should be assessed. (Lgr, 2011) The curriculum further states the importance of increased use of computers, not only to search information, but also as tools for learning.

Traditionally, students experience difficulties understanding the concepts of energy and power (Duit, 2007; Solomon, 1992; Wiser and Amin, 2001). Teaching materials have been developed to support teachers with this problem, both in terms of conceptual understanding, and connections to technology and sustainable development (Hobson, 2007; SEET, 2013; Connecticut Energy Education, 2013). Teaching strategies are often based on an investigative approach where the students have the chance to learn the subject content while they understand the situations where the content has value (Barab and Luheman, 2003). Although teaching materials have been developed and curriculums have changed, evaluations (for example PISA, 2009) show that Swedish students still have difficulties solving problems and applying their knowledge in science. Therefore, there is still a need for educational materials supporting teachers.

Teaching within design and product development has increased over the last few years due to its implementation in the curriculum. However, the research and development of practical teaching materials to assist the teaching and assessment of such topics is still needed. (Ritz and Martin, 2013, p. 391):

> "research needs included determining what knowledge and abilities that designing actually involves, criteria for evaluating novice designs, gaining conceptual knowledge through designing, and teacher trainee conceptions of design"

The teaching material described in this article, which among other things uses a computerized visualization product for examinations in the classroom and conceptual understanding, is based on successful teaching strategies within the energy field. Additionally, one part of the material focuses on the product development process.

2 THEORETICAL BACKGROUND

According to research in science and technology education, students understand the scientific concepts better when they are related to society and technology issues (Bennett et al., 2003). Course plans have also been developed to involve the contextualization and focus on sustainable development, which requires a change of teaching. Research has shown that an increased degree of exploration in different real contexts means that students increasingly appreciate the subject content (Barab and Luehmann, 2003). By teaching current social issues, such as climate change while teaching the traditional concepts of physics, Space (2007) describes how the concepts become more useful, relevant and interesting to the students. The students start to ask questions about discussions they have heard at home, seen on TV or read about in the newspaper. Space also describes student-centred discussions as successful as they develop an interest in physics concepts that did not exist in earlier education.

Research also shows that students become more engaged by project-based, student-centred and interdisciplinary education that is relevant to real life (Blumenfeld et al., 1996; Barab et al., 2000). Project-based teaching allows the students to work with authentic problems and ask their own relevant questions. This motivates them to find solutions they can explain and argue for (Blumenfeld et al., 1996).

Studies show that visualization in varying shapes makes it easier for students to understand abstract concepts, especially when they can study different phenomena and test hypotheses using visualization tools (Kozma et al., 1996). While working with the visualization tools, students should be given the opportunity to express their own thoughts and allow feedback from others. Additionally, a teacher should be able to follow the thought process of a student throughout the exercise. This should also occur in situations that do not include written or oral presentations (Lehtinen and Repo, 1996). If the visualization provides a common view that explains the problem, for example a screen shot that can be studied and referred to, research shows that such collaboration increases the meaningfulness (Anderson et al., 2000).

The teaching material developed within this project aims at providing the students with a visualization tool to explore and understand, among others, the concepts kilowatt-hour, power and energy, through reality-based examinations. By using the material the problems can be visualized and analysed, measurements can be critically examined, and various alternatives can be compared. According to researchers in science and technology education, students should be given the opportunity to plan and conduct their own investigations, make hypothesizes, search for information, and create models, so they can develop these skills. Also important are debates with peers and teachers, which give students the chance to use their knowledge and strategies to take a position and use well-formulated arguments. (emphasized by for example Linn et al., 2004).

The teaching material further allows the students to discover the process of developing a digital visualization tool. It lets them conduct reality-based design tasks in which they work practically, adopting professional roles. De Vries (2012) claims that project work with reality-based tasks, combined with computer modelling and visualization, works well to integrate the subjects of physics, technology, engineering science, and mathematics, and allows further insights into engineering.

This teaching material is based on the idea that, by starting from their own ideas and in interaction with the digital teaching material, students and teachers can together merge the ideas into coherent chains of thoughts and investigations (characteristics of a successful dialogic teaching, according to Alexander, 2004). The material also provides the students the opportunity to reflect upon their own choices and its consequences, which are important bases for understanding, in particular within technology (Pitt, 2006). The material is also based on the following specific strategies, favourable for learning according to Dylan Wiliam (2007): (1) learning objectives must be clearly defined, understood and shared by everybody, (2) discussions and tasks in the classroom need to show if and how learning in class works, (3) teachers must give students feedback that help them progress, (4) teachers should in different ways use classmates as resources to each other and (5) the teachers need to make students own their own learning, for example by self-evaluation and reflection.

3 MATERIAL DESCRIPTION

The teaching material r(e)flect aims at providing students with increased awareness and tools to reflect upon energy use and related behaviour. It is not a normative material, claiming right and wrong energy behaviour. The long-term goal is rather that students, by using r(e)flect, acquire enough knowledge about energy use and its consequences that they can make better informed choices regarding their own energy-related behaviour.

r(e)flect consists of four different parts, (1) the visualization tool (e)lVis which clarifies the concepts of energy and power and enables real-time experiments, (2) the product development process, (3) reality-based practical exercises and (4) the WattVett challenge that enables self-reflection upon energy behaviour. The new approach of this teaching material is twofold:

1. the ***combination*** of the four parts in the same material (there are other teaching materials focusing on one or two at the time)
2. all parts following the common theme of ***energy*** (where electricity visualization tools are a recurring issue in all four parts)

3.1 r(e)flect Box

The teaching material consists of a website (see Section 3.3.) and a physical material, see Figure 1. The packaging has been given an interesting and exciting design to attract both students and teachers.



Figure 1: Parts of the teaching material r(e)flect are packaged in an interesting and excitingly designed box.

When opening the lid, one finds a mirror with the minicomputer of the electricity visualization tool (see next section) mounted on it. The mirror both symbolizes the self-reflection and ”to look oneself in the eyes”. Apart from the computer-based visualization tool, the physical box also includes material for different activities connected to energy use, concepts used in physics and technology as well as developing skills in presentation and argumentation. The box can be borrowed by teachers free of charge.

3.2 Electricity Visualization Tool (e)lVis

The electricity visualization tool (e)lVis (Figure 2) is the main reason for a physical material. It aims at enabling experiments with electrical appliances from everyday life. It clarifies the concepts of energy and power by real-time representation of electricity use. Furthermore, electricity visualization tools are a common theme throughout the teaching material. In the material, the students can follow the product development process that led to a similar visualization product. The two solutions, built on similar systems of components, are compared. (e)lVis has been designed to stand long-term use in the school environment, while the other product suits workplace applications. This also provides the students with inspiration to their own project work.

(e)lVis is composed of a set of smart plugs (sensors), a minicomputer, and a knob. Although the hardware is made of commercial products, the software was custom developed for this project.

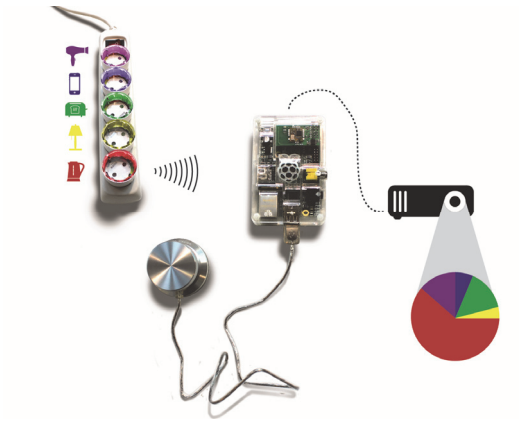


Figure 2: Illustration of a setup with the visualization tool (e)lVis. The tool comprises five smart plugs, a knob and a minicomputer. The minicomputer is connected to the classroom projector to show the measurement and result, i.e. a pie chart where each colour represents a device.

By locating one of the smart plugs between an electrical device and the power outlet the electricity use of appliances can be measured. Each plug is associated with a particular colour. The total amount of energy to be used within the experiment is set with the knob. As the measurement starts the electricity use of each connected device can be followed by its colour in a pie chart. The minicomputer is connected to the classroom

projector to show the real-time measurement and pie chart to the students on a screen or the wall. When the full amount of energy has been used, the proportion and number of watt-hours used by each appliance, along with the duration of the complete measurement will be shown, see Figure 2.

3.3 r(e)flect Website

All material included in the box, besides the visualization tool, can be downloaded from the website. This makes r(e)flect widely available to Swedish teachers. The website also provides instructional videos for teachers and engaging informational videos for the students. Throughout the website teachers can find hints, or how to sections, for how the material can be used and linked to the traditional classroom learning. The website is under development, but some screenshots from the current website are shown in Figure 3 and Figure 4.



Figure 3: Front page of r(e)flect website (under development).



Figure 4: Website page showing the four parts of the teaching material (under development).

3.4 WattVett Challenge

An important feature found on the website is the WattVett challenge. It has not yet been implemented, but the concept and how students are supposed to work with it is described in Figure 5.

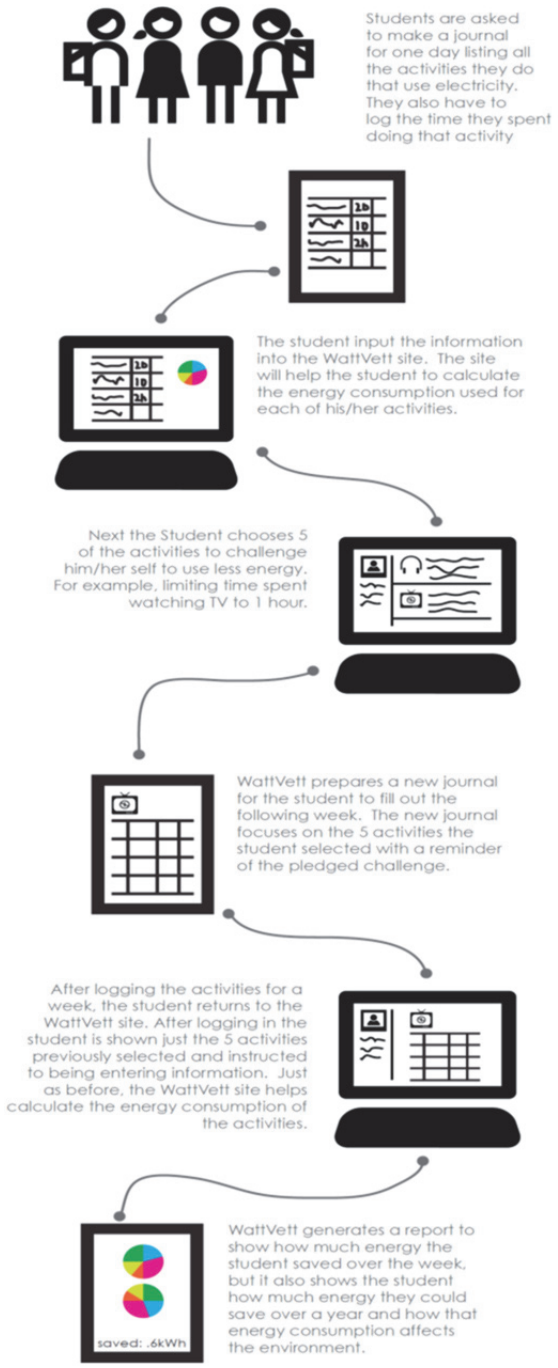


Figure 5: The students' work flow with the WattVett challenge.

In the WattVett challenge, students are invited to challenge themselves or other students and classes. The aim is to encourage self-reflection on the energy behaviour, but also to encourage an energy saving behaviour. Students report their energy use by

making an energy diary. Additionally, the self-reporting system can be used to follow up the energy saving potential connected to the teaching material.

4 PRELIMINARY RESULTS

The teaching material was evaluated in two different schools and with totally eight teachers. About 150 students tested different parts of the material. By using short questionnaires, observations and a few interviews we collected feedback and response on specific activities. All eight teachers appreciated the material because it generated interest from the students. Teachers commented in both interviews and questionnaires that the material had a "nice and professional" feel and that it contributes to an entrepreneurial approach, that it helps visualize the students' own electricity usage, and that it can contextualize physics and technology. Teachers appreciate that the activities are linked to curriculum and assessment criteria, and enables cross-curricular elements. All teachers further stress the importance of a teaching material that contains "everything", that clearly describes what needs to be prepared and sourced prior to an activity, e.g. current and relevant links to reports, websites, etc. and reproducible data, forms, etc. These resources are of great value because they are provided within the material.

Students aged 10-15 have used the material in the ordinary physics and technology education. In total, 50 students filled in a feedback questionnaire, another 50 students were observed and six students were interviewed. In all 50 questionnaires the students expressed excitement that "something unusual" happens with a device in the classroom and that they can see the measurement in progress and in real time. Teachers that were interviewed told us that the students can then correlate the measurement to the use of their own electrical appliances.

In the interviews, all eight teachers further expressed appreciation for the tool's measurement of a real situation that can be followed here and now, but also that the students could follow the electricity usage of different products. They can initially guess which one uses the most electricity and then receive a result directly. Subsequent discussion of the concept of power is relevant and highlights the need to consider how long a device is in use.

About 50 of the students in the trial were part of a debate, which proved to be an excellent activity in middle school, even though two teachers expressed concern with the activity in interviews. Observation of the debate did, however, highlight that some basic knowledge must be substantiated. Within the debate students must take a stand for different energy sources and use evidence-based arguments. The debates showed the importance for students to have knowledge about, for example, how electricity is generated by generators and solar applications. The debate also highlighted the importance of the teacher as moderator, guiding the discussion, spreading the word and picking up interesting threads.

About 50 students in lower secondary school tested the "electricity diary" within a homework assignment. The students first measured the electricity use of some "regular" devices of their own choice in the classroom. They reflected upon their everyday life and use of electrical appliances. As homework, they went home and noted their use of electricity during the day, talked to other family members, reflected on the effects on the electricity bills, and estimated the time and measurement of electricity use and costs. In interviews with six students, but also when observing discussions in the classroom, we found that the students, within their reflections, were given opportunities to reflect on aspects they would otherwise never have done.

Moreover, in interviews, two teachers said that the activities contributed to practical lessons for the students, such as realizing that the time aspect plays an important role for the electricity use. The exercises made the concepts power and watt visible to the students, and made them realize that both power and duration has to be taken into account when discussing total electricity use.

In interviews, all eight teachers expressed appreciation for teaching materials with the possibility to make power usage visible to students. They also emphasized the importance of a teaching material that exemplifies idea generation and product development, something that they lacked.

5 FUTURE WORK

We will continue developing the teaching material based on the initial evaluation. Twenty complete teaching kits have already been produced and distributed to schools.

In March 2014, we will let about 10 classes (i.e. around 250 students aged 10-15) and 10 teachers use and evaluate the teaching material. Students' perception of the material will be in demand, and their knowledge of energy will be tested after certain activities. Feedback will be collected through questionnaires, observations and interviews. The teachers' views will be sought through

questionnaires. Other students and teachers will also use the material during this period, but they will not be part of the formal evaluation. The evaluation will be based on methods described in Robson (2011).

At the end of the project, in 2015, both the web-based platform and the physical teaching material will be in their final design.

ACKNOWLEDGEMENTS

The project is funded by the Swedish Society for Nature Conservation (Naturskyddsföreningen) thanks to the sale of electricity labelled "good environmental choice".

REFERENCES

Alexander, R. J. (Red.). (2004). *Towards Dialogic Teaching: rethinking classroom talk.* North Yorkshire: Dialogos.

Anderson, A., Cheyne, W., Foot, H., Howe, C. Low, J. & Tolmie, A. (2000). Computer support for peer-based methodology tutorials. *Journal of Computer Assisted Learning, 16*, 41–53.

Barab, S.A., Hay, K. E., Barnett, M.G., & Keating, T. (2000). Virtual solar system project: Building understanding through model building. *Journal of Research in Science Teaching, 37(7*), 719-756.

Barab, S. A., & Luehmann, A. L. (2003). Building sustainable science curriculum: Acknowledging and accomodating local adaption. *Science Education, 87(4)*, 454-467.

Bennett, J., Hogarth, S., & Lubben, F. (2003). *A systematic review of the effects of contextbased and Science-Technology-Society (STS) approaches in the teaching of secondary science: Review summary.* University of York, UK.

Blumenfeld, P. C. et al. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher 25(8)*, 37 – 40.

Connecticut Energy Education (2013), The Institute for Sustainable Energy. Retrieved 2013 from Web site: http://www.ctenergyeducation.com/

Duit, R. (2007). Bibliography - STCSE (Students' and Teachers' Conceptions and Science Education). Kiel: IPN. Retrieved 2013 from Web site: http://www.ipn.unikiel.de/aktuell/stcse/stcse.html.

Hobson, A. (2007). Resource Letter PSEn-1: Physics and society: Energy. *American Journal of Physics. 75 (4)*, 294 - 308.

Kozma, R. B., Russell, J., Jones, T., Marx, N. & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. I: S. Vosniadou, E. De Corte, R. Glaser & H. Mandl (red.) *International Perspectives on the Design of Technology-Supported Learning Environments.* NJ: Erlbaum.

Lehtinen, E. & Repo, S. (1996). Activity, social interaction, and reflective abstraction: learning advanced mathematical concepts in computer-environment. I: S. Vosniadou, E. DeCorte, R. Glaser & H. Mandl (red.) *International perspectives on the psychological foundations of technology-based learning environments* (ss. 105–128). Hillsdale, NJ: Lawrence Erlbaum.

Lgr, 2011. Curriculum for the compulsory school, preschool class and the recreation centre 2011. http://www.skolverket.se/om-skolverket/andra-sprak-och-lattlast/in-english/publications.

Linn, M. C., Davis, E. A., & Bell, P. (2004). *Internet Environments for Science Education.* New Jersey: Lawrence Erlbaum Associates, Inc. Publischers.

PISA (2009) http://www.oecd.org/ Comments by Jacobsson, Oscarsson, Karlsson: retrieved 2013 from Web site: http://www.skolaochsamhalle.se/flode/skola/jacobsson oscarssonkarlsson-blir-svenska-elever-allt-samre-i-naturvetenskap/

Pitt, J. C. (2006). Human beings as artifacts. I J.R. Dakers (Red.), *Defining technology literacy: Towards an epistemological framework.* (s.133-144). New York: Palgrave Macmillan.

Ritz, J. & Martin, G., 2013. Research Needs for Technology Education: Highlights of a Preliminary Analysis. In *PATT27 Technology Education for the Future: A Play on Sustainability.* Editor P J Williams ISBN 978-0-9922497-2-4 New Zeeland: University of Waikato. Retrieved 2013 from Web site: http://www.iteea.org/Conference/PATT/PATT27/PATT27proceedingsNZDec2013.pdf.

Robson, Colin (2011). *Real World Research 3e.* ISBN 978-1-4051-82409. UK: John Wiley & Sons.

SEET. The Institute for Sustainable Energy SEET--Sustainable Energy Instructional Materials. Retrieved 2013 from Web site: https://ateec.org/professional-development/workshops-institutes/sustainable-energy-education-training-seet/next-workshop/

Solomon, J. (1992). *Getting to know about Energy in school and society.* London: Taylor & Francis Group.

Space, W. (2007). Climate Physics. Using basic physics concepts to teach about climate change. *The Science Teacher. 74* (6), 44 – 47.

Wiliam, D. (2007). *Research Brief: Five "Key Strategies" for Effective Formative Assessment.* Reston: The National Council of Teachers of Mathematics.

Wiser, M., & Amin, T. (2001). Is heat hot? Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and Instruction, 11,* 331-355.

Vries, M. J. de (2012), *'Teaching for Scientific and Technological Literacy: An International Comparison', in: Pfenning, U. and Renn. O. (Hrsg.), Wissenschafts- und Technikbildung auf dem Prüfstand.* Baden-Baden: Nomos Verlag, 93-110.

Supporting Creative Design Processes for the Support of Creative Mathematical Thinking

Capitalising on Cultivating Synergies between Math Education and Environmental Education

Chronis Kynigos[1,2] and Maria Daskolia[1,3]
[1]*Computer Technology Institute and Press "Diophantus" (CTI), 26-28 Mitropoleos str, 10563 Athens, Greece*
[2]*Educational Technology Lab, Department of Pedagogy, University of Athens, Athens, Greece*
[3]*Environmental Educational Lab, Department of Pedagogy, University of Athens, Athens, Greece*
kynigos@cti.gr, mdaskol@ppp.uoa.gr

Keywords: Mathematical Creativity, Social Creativity, Design of Educational Digital Resources, c-Book Technology, Communities of Interest, Environmental Education, Education for Sustainable Development, MC^2 Project.

Abstract: Mathematical creativity is acknowledged as a backbone lifelong competence necessary to be fostered in all students. However, this is not an easy task to accomplish, due not only to a lack of appropriate technologies enabling the creative design of digital educational resources for creative mathematical thinking (CMT), but also to the absence of insightful methodologies to support creative design processes of this kind among professionals. The Mathematical Creativity Squared (MC^2) project aims to address this twofold problem in the following ways: (a) by designing and developing a new genre of technological environment for the design of CMT resources, 'the c-book' environment, and (b) by adopting and further developing a methodology based on the generation of Communities of Interest (CoI) as a social milieu which will facilitate the creative design of CMT resources in collectives of educational designers stemming from diverse professional and educational domains. Especially with regards to the latter, the rationale is that CoIs will support synergies among designers with a math education background with others carrying a more socially-relevant educational orientation, such as environmental educators and educators for sustainable development, both on epistemological and pedagogical level, with the aim that their 'boundary crossing' interactions will positively effect social creativity in the design process for digital resources for CMT.

1 INTRODUCTION

1.1 On Creativity and Mathematical Creativity

Creativity is perceived as the backbone of the skills required for new jobs and as an essential ability to be developed in the context of lifelong learning (EC, 2008; 2011). It is recognised as a transversal skill needed to foster each of the eight key lifelong competences identified by the European Commission (EC, 2008) as particularly necessary for personal fulfilment and development, social inclusion, active citizenship and employment. Mathematical competence is among these core life-long competences identified by the European Commission, comprising the ability to develop and apply mathematical thinking to solve a range of problems in everyday situations and the application of this knowledge and methodology in response to perceived human wants or needs (EC, 2006). Developing students' creative potential in mathematical thinking has been also ranked among the most needed skills to develop for the 21st century in the States (National Academies of Science, 2007).

Nevertheless, although creativity is deemed as a tool for succeeding personal development and social empowerment and as an impetus for achieving professional innovation and economic change (Banaji et al., 2006), it still remains an elusive target to attain and a largely under-researched topic, especially in terms of whether and how it can be enhanced. The wide array of theories and perspectives on 'creativity' (e.g., Runco, 2007), coupled with the inherent vagueness of the concept (e.g., Kampylis and Valtanen, 2010), its context-specificity (Amabile, 1983; Csikszentmihalyi, 1996), and the inadequacy of most traditional educational

systems and processes, set several burdens in the fostering of creativity.

This is particularly true with mathematical creativity. The lack of an accepted definition among researchers and practitioners (Mann, 2006) on the one hand, and educational reality which systematically deemphasize creative math problem-solving ability on the other cannot but hinder any concerted activity in the field (Mann, 2009; Sriraman, 2005). Even though almost any academic perspective on math education has been promoting mathematical creativity (Menghini et al., 2008), well-meant intentions are not enough for developing math learning through creativity and vice versa. Endeavours of this kind are acknowledged to be closely associated with putting forth changes in educational systems, processes and outcomes; taking full advantage of the potential of information and communication technologies; and working towards materialising creative ideas into concrete, new and more effective, products and services (EC, 2008; Ferrari et al., 2009).

1.2 Fostering Mathematical Creativity through Digital Technologies

In terms of technologies, twenty or more years ago, emergent approaches to how digital media could be used to foster mathematical thinking were almost exclusively visionary and about using technology as an expressive medium for cutting edge pedagogy involving engagement with mathematical thinking and meaning making (Hoyles and Noss, 2003). Despite the existence of some excellent digital tools affording dynamic manipulation, interconnected representations (including mathematical formalism), simulations of phenomena and situations embedding mathematical rules, visualisations of data representations and handling of probability, their uses in education were frequently instrumented to a large extent towards contexts of traditional lecturing and demonstration of exercise-type activities (Hennessy et al., 2005; Ruthven, 2008). Some technological advances of the past few years have even further skewed attention back to uses of digital media to support traditional routine learning of mathematics through the assistance of learning management systems, portals, video and intelligence in tutoring systems to that effect (e.g., the Kahn Academy).

However, especially during the last decade there is a growing interest in Europe in the role and use of appropriate digital technologies as the necessary broad-based strategies to bring forth the envisioned shifts in terms of creativity and innovation in education (Ferrari et al., 2009). This rests to be explored with regards to the teaching and learning of any knowledge domain, including mathematics. Nevertheless, whether digital technologies can act as real enablers of creative mathematical thinking has to be combined with other processes, mechanisms and tools of school education, among which the design and use of appropriate educational resources.

1.3 Creativity in the Design of Educational Resources for CMT

This last point brings forth an inter-related and equally challenging issue, that of the design of appropriate educational resources for creative mathematical thinking (CMT), and its relation to creativity. Taking into consideration that any kind of design process is inseparably connected to creativity (Taura and Nagai, 2010), the task of designing educational resources for CMT becomes thus a 'squared' creativity issue.

Design creativity is recognised as a multidimensional concept, that can be identified, among others, in the processes or the outcomes of the design activity, or in the context within which such an activity is embedded (Gero, 2010). This is also the case with the design of educational resources (and instructional design in general), which although not overly acknowledged and studied as a design discipline compared to other domains (i.e., software design or architecture), creativity has recently also being acknowledged as a ‘built-in’ dimension of it (Clinton and Hokanson, 2012).

The 'social' component in various collectives of design professionals and its role in enhancing both the individuals' creativity and the creative capacity of the group in addressing complex design problems is also a relatively new focus of research. 'Social creativity' actually explores the social and technical environment within which participatory design processes take place (Fischer, 2001; 2011), when specialists from different domains coordinate their efforts to achieve a common design goal. Social creativity arises from the synthesis and synergy of the different perspectives towards a complex design task of shared interest.

Designing for digital educational resources for CMT can be therefore viewed as a squared creativity challenge, since it requires not only to define mathematical creativity but also to situate the design process itself within a socio-technical environment that can boost educational designers' creative

potential. However, both challenges to be dealt with appropriately need some paradigm shifts in terms of thinking of, learning and designing pedagogical materials for mathematics, already represented in some debates in math education.

1.4 Boosting Math Creativity through Synergy with Other Domains: the Case of Environmental Education

For example, a key debate in the field is the one questioning the traditional paradigm of focusing exclusively on abstract mathematical concepts and problems as isolated from broader phenomena and social contexts in which they could become more relevant, and without highlighting the relevance and value of this knowledge to the students' life and society in general. Within this paradigm, math education seems to reproduce the false myth of an objective and value-free discipline, alienated from current reality. Criticisms of this genre are supported by evidence from many countries indicating that school curricula give preponderant emphasis on foundationalist approaches of learning about mathematics and on pedagogies advancing the transmission of dry knowledge, as if all students were being prepared to become the big mathematicians of tomorrow. However, there are serious doubts about whether such curricula can succeed to trigger in students any meaningfully creative engagements with maths.

A suggestion so that math education explores new creativity potentials and meaningfulness to a wider range of students is to 'bridge' it with educational domains which are more socially-oriented and centred to real-life problems, such is the case of Environmental Education (EE) and Education for Sustainable Development (ESD) (Kynigos et al., 2013). Although several scholars have stressed the opportunities for developing a beneficial relationship between science education and EE/ESD (such as Gough 2002, 2007; Sjøberg and Schreiner, 2005, etc), no relative bridging has been proposed to motivate students to get more actively involved with math concepts and processes by identifying the 'mathematics' hidden inside some of the most challenging current socio-scientific and sustainability issues.

Especially mathematical problem-posing and problem-solving which are indicated as appropriate learning formats to allow math creativity to emerge, are proposed to exemplify some characteristics with respect to the 'problematic situation'. For example, for a situation to be amenable to creative (mathematical) problem-posing and -solving (Torp and Sage, 2002) it:

- should be fairly ill-structured and messy
- may change with the addition of new information
- is not solved easily or with a specific formula
- does not result in one right answer.

These characteristics are directly applicable to most environmental and sustainability issues, which are by nature ill-defined, complex, controversial, value-laden and require the application of various perspectives to grasp them more thoroughly (Daskolia and Kynigos, 2012). However, it is this messiness and complexity that make them 'good' examples of creativity-triggering problems and amenable to be treated in learning situations that foster the students' creative thinking. It also turns them into "boundary objects" (Star and Griesemer, 1989), that is entities that can interpreted and employed by more that one groups or communities in ways that make sense to them.

At the same time, dealing with issues of these characteristics at a pedagogical level provides many opportunities for teachers and learners to get engaged in dialogical forms of meaning-construction and perspective-sharing and to expand the "boundaries" of their knowing of and being in the world both inside and across the realms of their discipline (Daskolia and Kynigos, ibid). This is what some would also identify as a creative appropriation of subject-matter and pedagogical knowledge.

2 THE PROJECT RATIONALE

For mathematical creativity to be fuelled new designs are needed, new ways of thinking and learning about mathematics through synergy with other educational domains, such as those of EE and ESD, and the support of learners' engagement with creative mathematical thinking in collectives using dynamic digital media. Designers of educational digital resources (either professionals or teachers) have therefore to look for and benefit from new designs that would assist them to explore, identify and creatively produce new educational resources and tools with a potential to stimulate creative ways of mathematical thinking.

This is not an easy task to accomplish, due not only to a lack of appropriate technologies enabling the creative design of digital educational resources for creative mathematical thinking (CMT), but also to the absence of insightful methodologies to support creative design processes of this kind. The MC^2

project aims to address this twofold problem in the following ways: (a) by designing and developing a new genre of technological environment for the design of CMT resources, that is an authorable e-book we call **'the c-book'** (c for creative), and (b) by adopting and further developing a particular methodology based on the generation of **Communities of Interest** (CoI) (Fischer, 2001; 2011) as a social environment that would facilitate the creative design of CMT resources in collectives of professionals. The rationale is that by providing educational designers with an appropriate socio-technical environment, they will have more opportunities for richer creative design processes.

The objective of the MC^2 project is thus to build both a technology and a social milieu that will jointly and interactively enhance creativity among educational designers in the design of digital educational resources fostering creative mathematical thinking in learners. Our approach addresses *social creativity in design* through specially formed Communities of Interest (CoIs) which will act as agents of potential cultures of participation among professional designers and teachers from different/ diverse educational domains (not only from the math education sector but also from environmental education) and will lead to the development of innovative frames of addressing the design of educational resources for creative mathematical thinking.

The MC^2 project places at the centre of its concerted action the design and development of ***the c-book,*** an integrated and beyond the state-of-the-art digital system which will facilitate all foreseen creative design processes and practices *and* will provide support for creative mathematical thinking. The main characteristics of this system will be:

- an authorable data-analytics engine and a graphical interface providing professionals from diverse creative industries involved in the design of digital resources for CMT (publishers, developers, researchers, school educators from diverse educational domains) with the ability of customizing the kinds of information they need for assessing end-user creativity and the kinds of automated reactions they wish the tool to provide to end-users' activity.
- an authorable dynamic e-book infrastructure to be used by pedagogical design professionals for collaborative design of CMT resources and by students to create their personalized versions of such resources.
- a set of dynamic, exploratory and constructionist digital tools integrated in the e-book infrastructure designed to foster creativity in students' mathematical expression, investigation and meaning generation.

The design of the c-book environment will require advancements not only in terms of technology but also in the design processes. MC^2 project will develop, operationalize and document methods for the emergence of creativity in the design process based on **'middle c'** social creativity in CoI embedding diverse actors and norms (Moran, 2010; Hämäläinen and Vähäsantanen, 2011; Sonnenburg, 2004). In effect, it will address and promote the creativity that emerges when communities share common goals and well-designed tools. The project partners have the capacity and experience to act as catalysts for systemic change in cultures characterizing current school and workplace practices. By situating the CoI in large-scale systemic national initiatives MC^2 will test how the process and productions of collaborative creative design for CMT can empower professionals (and teachers) from diverse (disciplinary and/or educational) domains to collaborate and create new educational designs. At the scientific level, the project will contribute to our knowledge of methods and processes in using digital media to enhance creativity in interdisciplinary CoIs. The emergence of creativity is perceived as 'a system' involving 'collectives with tools' characterized as 'cultures of participation' (Fischer, 2011). This project will advance our knowledge by integrating two constructs to generate and support such cultures for professional design, 'documentational approach' (Gueudet and Trouche, 2012) and the 'boundary crossing with digital media as boundary objects' approach (Akkerman and Bakker, 2011; Kynigos & Kalogeria, 2012). The technology will play a fundamental role by transforming an otherwise costly endeavour to an efficient knowledge co-construction process through both the collaborative and authoring features of the c-book which will be useable in the six languages of the partnership. This will provide stakeholders with the means to both advance and share creative intelligent mathematical artifacts and by thus to enable a broader reach. Accordingly, the impact of MC^2 will be long-lasting as the malleability of the creative widgets and the authorable intelligent feedback will create the foundations for subsequent evidence-based creative designs beyond the life of the project.

3 NEXT STEPS

Our position is that digital media designed to

support practices and productions of added pedagogical value fall short of target if they address only the conceptual/ cognitive/ performance aspects of the learning process. There is a need to consider how to directly support future citizens in their social and work capacity. Mathematical creativity is viewed as a backbone ability and disposition for students, which can be cultivated and taught as a social phenomenon per se but also through teaching and learning about more socially-related and focused issues (i.e., current environmental and sustainability issues) and be widened up to all students. Social creativity amongst designers of educational materials with an explicit aim to support such a creative potential in learners is a core competence which nevertheless needs some appropriate frames to be identified, studied and boosted. Technologies supporting the design and deployment of educational materials and digital media for enhancing the students' creative mathematical thinking have so far been too much influenced by more traditional paradigms of thinking where such media have the role of unquestionably accredited tools to be employed by the students. New genres of technologies, such as the proposed c-book system, will help the questioning and development of the process of design and the kinds of media and accompanying paradigms which may become available to students in the future. Within this rationale, a c-book unit can be a medium containing questionable artifacts, statements, arguments for students to change, improve, negotiate over, re-construct; something for the students to re-draft, to shape or make on their own. They are viewed as objects for students to argue over and to collectively 're-write'. How does this will consequently change the designers' perspectives, methodologies and ways of thinking, collaborating and working? Our proposed c-book technology and supporting methodology aim to re-address how to design educational resources for creative mathematical thinking by turning into a creative learning asset the crossing of various pre-established boundaries across many inter-related or not until recently related fields.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 610467 - project “M C Squared”. This publication reflects only the authors’ views and Union is not liable for any use that may be made of the information contained therein.

REFERENCES

Akkerman, S. F., & Bakker, A., 2011. Boundary Crossing and Boundary Objects. *Review of Educational Research, 81*(2), 132-169.

Amabile, T. M., 1983. *The social psychology of creativity*. New York: Springer-Verlag.

Banaji, S., Burn, A., and Buckingham, D., 2006. *Rhetorics of creativity: a review of the literature*. Arts Council England, Centre for the Study of Children, Youth and Media, Institute of Education, University of London, London, UK.

Clinton, G., Hokanson, B., 2012. Creativity in the training and practice of instructional designers: the Design/Creativity Loops model. *Educational Technology, Research and Development, 60*, 111–130. DOI:10.1007/s11423-011-9216-3.

Csikszentmihalyi, M., 1996. *Creativity: Flow and the Psychology of Discovery and Invention*. Harper Collins, USA.

Daskolia, M., Kynigos, C., 2012. Applying a Constructionist Frame to Learning about Sustainability. *Creative Education*, *3*, *special issue: 'Higher Education'*, 818-823. Published Online in SciRes (http://www.SciRP.org/journal/ce), DOI:10. 4236/ce.2012, Scientific Research.

EC, 2011. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A Single Market for Intellectual Property Rights Boosting creativity and innovation to provide economic growth, high quality jobs and first class products and services in Europe, http://ec.europa.eu/internal_market/copyright/docs/ipr_strategy/COM_2011_287_en.pdf.

EC, 2008. Lifelong Learning for Creativity and Innovation. A Background Paper: Slovenian EU Council Presidency, http://www.sac.smm.lt/images/12%20Vertimas%20SAC%20Creativity%20and%20innovation%20%20SI%20Presidency%20paper%20anglu%20k.pdf.

EC, 2006. Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning, *Official Journal L 394 of 30.12.2006, 10–18,* http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:EN:PDF.

Ferrari, A., Cachia, R. and Punie, Y., 2009. *Innovation and Creativity in Education and Training in the EU Member States: Fostering Creative Learning and Supporting Innovative Teaching. Literature review on Innovation and Creativity in E&T in the EU Member States (ICEAC).* European Commission - Joint Research Centre Institute for Prospective Technological Studies, http://ipts.jrc.ec.europa.eu/

publications/pub.cfm?id=2700.

Fischer, G., 2011. Social Creativity: Exploiting the Power of Cultures of Participation, Proceedings of SKG2011. In *7th International Conference on Semantics, Knowledge and Grids* (pp. 1-8). Beijing, China.

Fischer, G., 2001. Communities of Interest: Learning through the Interaction of Multiple Knowledge Systems. In *Proceedings of the 24th IRIS Conference* S. Bjornestad, R. Moe, A. Morch, A. Opdahl (Eds.) (pp. 1-14). August 2001, Ulvik, Department of Information Science, Bergen, Norway.

Gero, J. S., 2010. Future directions for design creativity research. In T. Taura & Y. Nagai (Eds), *Design Creativity 2010* (pp.15-22), Springer.

Gueudet, G., and Trouche, L., 2012. Communities, documents and professional geneses: interrelated stories. In G. Gueudet, B. Pepin, & L. Trouche (Eds.), *Mathematics curriculum material and teacher documentation: from textbooks to lived resources* (pp. 305-322). New York: Springer.

Gough, A., 2007. Beyond convergence: reconstructing science/ environmental education for mutual benefit. Keynote address at the *European Research in Science Education Association (ESERA) Conference*, Malmo, Sweden 25-28 August 2007.

Gough, A., 2002. Mutualism: A different agenda for environmental and science education. *International Journal of Science Education, 24*(11), 1201–1215.

Hämäläinen, R. & Vähäsantanen, K., 2011. Theoretical and pedagogical perspectives on orchestrating creativity and collaborative learning. *Educational Research Review, 6*(3), 169–184. http://dx.doi.org/10.1016/j.edurev.2011.08.001.

Hennessy, S., Deaney, R., Ruthven, K., 2005. Emerging teacher strategies for mediating technology-integrated instructional conversations: A socio-cultural perspective. *The Curriculum Journal, 16*(3), 265-292.

Hoyles, C., Noss, R., 2003. What can digital technologies take from and bring to research in mathematics education? In A. J. Bishop, M.A. Clements, C. Keitel, J. Kilpatrick and F.K.S. Leung (Eds.), *Second International Handbook of Mathematics Education*. Dordrecht: Kluwer Academic Publishers.

National Academies of Science, 2007. *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: National Academies Press.

Kampylis, P., and Valtanen, J., 2010. Redefining creativity—Analyzing definitions, collocations and consequences. *Journal of Creative Be-havior, 44,* 191-214. DOI:10.1002/j.2162-6057.2010.tb01333.x.

Kynigos, C., Daskolia, M., Smyrnaiou, Z., 2013. Empowering teachers in challenging times for science and environmental education: Uses for scenarios and microworlds as boundary objects. *Contemporary Issues in Education*, *3*(1), 41-65.

Kynigos, C. Kalogeria, E., 2012. Boundary Objects for in service Mathematics Teacher Education: the case of Scenarios and Half-baked Microworlds, Special Issue in *Online Mathematics Education*, *The International Journal of Mathematics Education - ZDM, Springer Verlag, 44*, 733–745. DOI: 10.1007/s11858-012-0455-5.

Mann, E., 2009. The search for mathematical creativity: Identifying creative potential in middle school students. *Creativity Research Journal, 21*, 338–348.

Mann, E., 2006. Creativity: The essence of mathematics. *Journal for the Education of the Gifted, 30*, 236–230.

Menghini, M., Furinghetti, F., Giacardi, L., and Arzarello, F. (Eds.), 2008. *The first century of the International Commission on Mathematical Instruction (1908 - 2008). Reflecting and shaping the world of mathematics education*. Roma: Collana Scienze e filosofia.

Moran, S., 2010. Creativity in school. In K. Littleton, C. Woods, & J. K. Staarman (Eds.), *International handbook of psychology in education* (pp. 319–359). Bingley, UK: Emerald Group Publishing Limited.

Runco, M. A. (Ed.), 2007. *Creativity - Theories and themes: research, development, and practice*. Amsterdam: Elsevier Academic Press.

Ruthven, K., 2008. Mathematical technologies as a vehicle for intuition and experiment: A foundational theme of the International Commission on Mathematical Instruction, and a continuing preoccupation. *International Journal for the History of Mathematics Education, 3*(2), 91-102.

Sjøberg, S., & Schreiner, C., 2005. Young people and science: Attitudes, values and priorities. Evidence from the ROSE project. In *EU Science and Society Forum*, Brussels, 8-11 March.

Sonnenburg, S., 2004. Creativity in communication: A theoretical framework for collaborative product creation. *Creativity and Innovation Management, 13*(4), 254–262.

Sriraman, B., 2005. Are giftedness and creativity synonyms in mathematics? *Journal of Secondary Gifted Education, 17*(1), 20–36.

Star, S. L., & Griesemer, J. R., 1989. Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science*, *19*, 387-420.

Taura, T., Nagai, Y. (Eds.), 2010. *Design Creativity*, Springer.

Torp, L., & Sage, S. 2002. *Problems as possibilities: Problem-based learning for K-16 education*. Alexandria, VA: Association for Supervision and Curriculum Development, 2nd edition.

Domain Applications and Case Studies

FULL PAPER

Fostering Information Literacy in German Psychology Students Using a Blended Learning Approach

Nikolas Leichner, Johannes Peter, Anne-Kathrin Mayer and Günter Krampen
Leibniz Institute for Psychology Information, Universitätsring 15, 54296 Trier, Germany
{leichner, peter, mayer, krampen}@zpid.de

Keywords: Information Literacy, Blended Learning, Distance Education, Psychology, College Students.

Abstract: This paper reports about the experimental evaluation of a blended learning course for information literacy tailored to the needs of undergraduate Psychology students. The course consists of three modules delivered online and two classroom seminars; the syllabus includes scholarly information resources, ways to obtain literature, and criteria to evaluate publications. For evaluation purposes, a multimethod approach was used: The students completed an information literacy knowledge test and three standardized information search tasks (ordered by ascending difficulty) before and after taking the course. A sample of $N = 67$ undergraduate Psychology students ($n = 37$ experimental group, $n = 30$ waiting control group) participated in the course. As it was expected, students' knowledge test scores as well as performance in the search tasks improved markedly during the course. Results are discussed with regard to the soundness of the evaluation criteria used and to further development of the course.

1 INTRODUCTION

The term Information literacy is used to describe the ability to realize when there is a need for information, and the ability to identify, locate, and evaluate additional information which is required to meet this need (National Forum on Information Literacy n. d.; American Library Association 1989). Against the background that advances in information technology lead to a growing number of information resources, information literacy can be considered a "basic skills set" (Eisenberg 2008).

As information literacy is of importance in nearly all circumstances (e.g. Eisenberg 2008), a clear definition of information literacy has to be provided first of all. Hence, we limit our research to information literacy in higher education, especially in Psychology. Our definition of information literacy is based on the ACRL Psychology information literacy standards (Association of College and Research Libraries 2010), because this framework includes detailed performance indicators. This definition includes four standards of information literacy:

1) Determining the nature and amount of information needed: exemplary performance indicator: "understands basic research methods and scholarly communication patterns in psychology necessary to select relevant resources";

(2) Assessing information effectively and efficiently: exemplary performance indicator: "selects the most appropriate sources for accessing the needed information";

(3) Evaluating information and incorporating information into one's knowledge system: exemplary performance indicator: "compares new information with prior knowledge to determine its value, contradictions, or other unique characteristics";

(4) Using the information effectively to accomplish a specific purpose; exemplary performance indicator: "applies new and prior information to the planning and creation of a particular project, paper, or presentation".

We decided to focus on standards one to three, as our course was devoted primarily to the improvement of information seeking skills. Skills related to standard four are part of the curriculum of academic writing courses at German universities; consequently, they should not be part of our course. With regard to our curriculum, we expected the participants inter alia to be (better) able to understand scholarly communication patterns, distinguish between several research methods (e.g. meta-analysis and empirical study), to select the most appropriate resource for their information need and to evaluate the literature found.

There is indication that incoming students are not

sufficiently information literate (Smith et al. 2013) and that students do not become information literate during the course of their studies (Warwick et al. 2009). To address this need, almost every university library in Germany provides information literacy courses that complement the academic writing courses offered by university departments. However, most of these courses are two-hour events covering only a facet of academic information literacy (e.g. the use of bibliographic databases). Due to limited time, these courses offer few possibilities to practice information seeking, or to ask questions and discuss matters. As there is no standardized test to assess information literacy for German speaking populations, in most cases, evaluation efforts are based on feedback provided by participants. Finally, most courses are not tailored to specific disciplines, or fields (e.g. economics, social sciences). This is problematic, as scholarly communication patterns and information resources differ between disciplines. Course content should therefore be adapted to the needs of Psychology students (e.g. Thaxton et al. 2004).

The aim of this is study is to create a blended learning approach to teach information literacy to undergraduate students in Psychology. At the start, the course will exist alongside the courses offered by university libraries. However; we hope that our materials (at least some of them) will later on be used by university libraries to offer courses tailored to Psychology students. An important component of our effort is carrying out an evaluation study based on performance indicators using a control group.

The main reason for choosing a blended learning approach was to give participants the chance to work on the online materials adapted to their individual schedule. This is particularly important as students are often pressed for time. However, there should also be traditional classroom teaching as online learning alone seems to be fraught with higher dropout rates (Carr 2000). There is research indicating that blended learning can reduce dropout rates (López-Pérez et al. 2011) and is more effective than traditional classroom teaching or online learning (Clardy 2009).

2 OUTLINE OF THE COURSE

2.1 Content

The content of the course was determinated based on the Psychology specific information literacy standards provided by the ACRL (Association of College and Research Libraries 2010) and based on our own considerations. The target group were undergraduate students, so the content was mainly basic information about

- scholarly communication patterns in Psychology and common publication types (e.g. empirical article, review article, edited books);
- different information resources (inter alia bibliographic databases, internet resources) and their advantages and disadvantages;
- appropriate use of these resources (e.g. understanding of the thesaurus and of Boolean operators);
- inclusion of resources provided by related disciplines (e.g. PubMed, ERIC) in case the topic is of interdisciplinary nature;
- options for the acquisition of literature (e.g. use of electronic journal subscriptions, the local library catalogue, or interlibrary loan);
- criteria for the selecting publications beyond their content, e.g. Journal Impact Factor.

2.2 Structure

As mentioned before, the course combined online and traditional classroom teaching. In total, there were three modules to be completed online and two classroom seminars. We expected that completing the online materials would take up to four hours; both classroom seminars were designed to take 90 minutes each. The course was scheduled to be completed within two weeks, what seems a reasonable workload for college students.

The concept of the course envisaged that most of the knowledge should be imparted by the online materials, while the main purpose of the classroom seminars was to provide an opportunity to solve information problems under the guidance of the instructor and to ask questions. For this reason, participants had to complete certain online materials before attending the related classroom seminar.

To present an example, online modules 1 and 2 were related to the first classroom seminar. These modules dealt mainly with scholarly communication patterns, information sources (and their functions), as well as the acquisition of literature. A central element of the related first seminar was the task to find scientific literature on the question how distractions impact car driving performance. At the beginning of the seminar, the task was presented to the participants. Then, the task was split into steps (determining the information need, finding search terms, conducting the search, selecting literature) and participants worked on the steps either individu-

ally, or in small groups. The instructor was available in case questions arose. After completing each step, one of the participants (one group, respectively) had to present his/her outcome to the other participants and the outcome was discussed.

All online modules were provided via the e-learning platform "Moodle". Most of the content was presented using short passages of text which were enriched by illustrations or screenshots of the relevant computer programs. This content was provided using lessons, or pages inside of Moodle. The materials also included several videos. At the end of every section of the course, short quizzes were provided, so that the participants could apply their knowledge right after learning.

3 EMPIRICAL STUDY

3.1 Instruments

In most cases, information literacy is assessed using knowledge tests consisting of multiple choice items (e.g. Project SAILS 2013; Center for Assessment and Research Studies 2013; Noe and Bishop 2005). These tests have been shown to provide a reliable, valid and economic way of measuring information literacy. However, as information literacy is a complex ability, it can be doubted whether a knowledge test can assess information literacy comprehensively. For instance, appropriate information seeking behavior is an elementary part of being information literate (Timmers and Glas 2010) and the assessment of information seeking behavior requires observation (e.g. Julien and Barker 2009), or self-reports (e.g. Timmers and Glas 2010).

Besides, several authors argue that competencies should be assessed using real-life tasks instead of knowledge tests (Shavelson 2010; McClelland 1973).

For these reasons, we decided to use a multi-method approach consisting of two standardized tests which were applied in a laboratory setting: a knowledge test and information search tasks.

The knowledge test consisted of 35 multiple-choice items and had previously been developed by our research group. When developing the items, we relied on Standards 1 to 3 of the aforementioned information literacy definition (Association of College and Research Libraries 2010). A sample item is:

Which differences exist between Internet search engines (e.g. Google Scholar) and bibliographic databases?

a. *bibliographic databases usually have a thesaurus search*
b. *Boolean operators can only be used with bibliographic databases*
c. *the order of items on the results page is not affected by the number of clicks on each item*

The test had been used in a previous study with a sample of $N = 184$ participants who had completed the test online. In this study, an acceptable internal consistency of the test of Cronbach's Alpha $\alpha = 0.49$ was found. Furthermore, it was found that Master level-students scored significantly higher than undergraduate students in their first and second year. These results can be considered an indication of the validity of the test.

The information search tasks are based on a taxonomy of tasks from which instances of tasks can be derived. When reviewing the literature on information search tasks, we found that the existing taxonomies are of a descriptive nature and do not provide indications for the difficulty of a certain task type (e.g. Kim 2009). Another problem with these taxonomies was that they had been developed to classify non-scholarly search tasks in electronic resources. There are several differences between academic and non-scholarly searches; the most important one might be the use of bibliographic databases (e.g. PsycINFO) instead of internet search engines (e.g. Google, Yahoo!).

For these reasons, we decided to develop a task taxonomy specifically for academic information search tasks in Psychology. The taxonomy provides three types of information search tasks differing in their difficulty. To be more precise, the tasks differ in the abilities and competencies required to solve the task. The taxonomy is designed in a way that abilities required to solve tasks of the first type are also required to solve tasks of type 2. However, solving tasks of type 2 requires additional competencies, as do tasks of type 3. The taxonomy can be used to develop several tasks of the same structure and difficulty which can be used for assessing information literacy. For illustration purposes, a type 2 task (medium difficulty) is provided:

Are there meta-analyses published after 2005 investigating "risk factors" for the development of a "Posttraumatic stress disorder"? If possible, indicate two publications.

To solve this task, the participant has to under-stand the keyword search function in a bibliographic database, and needs an understanding of Boolean operators and complex filter functions in bibliographic databases to find publications using a certain meth-

odology (e.g. meta-analysis). Type 1 tasks are easier as they do not require an understanding of Boolean operators and complex filter functions. Type 3 tasks are more difficult, as they additionally require the participant to identify appropriate search terms before conducting the search. To score the tasks, rubrics for scoring the search task outcome (which publications were found) and the procedure applied by the students when completing the tasks were created. In line with the rubric for scoring the outcome, scores were awarded depending on how close the publications found come to the requirements mentioned in the task description (e.g. thematic focus of the publication, publication date). As stated in the rubric for scoring the procedure, scores were awarded for working on the tasks in an efficient and information literate way as defined by the information literacy standards (Association of College and Research Libraries 2010). For example, for a type 2 task, the maximum number of procedure scores was rewarded if the participant solved the task using bibliographic databases, used Boolean operators to combine two search terms and limited the results using the corresponding functions of the database.

3.2 Method

3.2.1 Participants

The sample consisted of $N = 67$ undergraduate Psychology students who took the course. Out of these students, 34 were first year students, while 33 were second year students. The average age was 21.67 ($SD = 2.38$). Participants had agreed to additionally participate in three data collection sessions for which they were compensated. Participants were randomly assigned to one of two groups; group 1 (experimental group) consisted of $n = 37$ participants, group 2 (waiting control group) of $n = 30$.

3.2.2 Procedure

The duration of the evaluation study was four weeks, while the actual course took only two weeks. Data collection 1 took place right at the start of the evaluation study. Subsequently, group 1 participated in the course, while group 2 served as a waiting control group. Two weeks later, when the course was completed for group 1, data collection 2 took place. After that, group 2 participated in the course. The final data collection took place after all participants had completed the course.

The data collections were scheduled in the computer lab of Trier University. Students were tested in groups of 15 to 22 participants under the supervision of two experimenters. They first completed three information search tasks (one task of each type), followed by two questionnaires concerning epistemological beliefs, and the information literacy knowledge test.

To complete the information search tasks, the participants could use all resources available from these computers (access to the internet, to bibliographic databases, and to the online library catalogue). The information search tasks were presented ordered by difficulty. Participants were required to record the publications found by using input boxes that were provided by the software used. After the completion of every task, the participants had to answer several questions concerning the procedure of their search. These data were the basis for scoring the procedure. As it would be beyond the scope of this paper, the results concerning the epistemological beliefs questionnaires are not reported.

3.2.3 Hypotheses

With regard to the search tasks, we expected that those tasks requiring more abilities should be more difficult; in this case, we expected type 3 tasks to be the most difficult followed by type 2 tasks, which, in turn, should be more difficult than type 1 tasks (Hypothesis 1).

As explained above, there were three variables designed to assess information literacy: outcome and procedure of the search tasks, and the knowledge test. We expected to find significant correlations among these instruments. As the knowledge test had already been tried and tested in a different study, finding correlations between the test and the search tasks would corroborate the status of the search tasks as indicator of information literacy (Hypothesis 2).

Furthermore, we expected that participants would score higher on all instruments after participating in the course. Specifically, group 1 should outperform group 2 at data collection 2. At the final data collection, there should not be any difference between the groups (Hypothesis 3).

3.3 Results

Before the course could be evaluated, the information search tasks had to be scored independently by two raters. The inter-rater-reliability (correlation between the scores awarded by the two raters) was in the range from $r = 0.62$ to $r = 0.92$; most correlations were above $r = 0.70$. In those cases where the scores differed, the raters agreed on one solution which was used for the analyses.

The first hypothesis to be examined was whether the expected order of task difficulties could be verified empirically. Data from data collection 1 is presented in table 1. As can be seen, tasks of type 3 were more difficult than tasks of type 2, which, in turn, were more difficult than type 1 tasks. The table shows the percentage of the maximum score for the different tasks types.

Table 1: Percentage of maximum score for search task outcome and procedure at data collection 1.

task type	outcome	procedure
1	77%	55%
2	50%	46%
3	32%	36%

For the following analyses, the outcome and procedure scores of each data collection were summed up separately, so that 2 scores for each data collection resulted. These scores were scaled, in order to restrict their range from 0 to 1 and are presented in table 2. Before using these scores for evaluating the course, we determined whether there were differences between the two groups of participants before the course started. Our analysis revealed that there were no differences between the groups, neither on the outcome scores (t[65] = 1.34, n.s.), nor on the procedure scores (t[65] = 1.23, n.s.). Furthermore, the two groups did not differ in their performance on the knowledge test (t[65] = 0.78, n.s.). Scores on the information literacy knowledge test were also scaled to restrict their range from 0 to 1, and can be found in table 3.

To examine the second hypothesis, correlations between the scores on the knowledge test and the two search task variables were computed using data from data collection 1. It was decided to analyze data from data collection 1 only, as the performance at the following data collections reflects to a great extent how much the participants have benefited from the course, so the results might be distorted.

The outcome and procedure scores of the search tasks correlated significantly (r = 0.22, p < 0.05), even though the correlation was weak. Both scores also correlated significantly with the performance on the knowledge test (for the outcome scores r = 0.29, p < 0.01, and the procedure scores r = 0.48, p < 0.01, both one-tailed).

To evaluate the course (Hypothesis 3), the three information literacy performance indicators were analyzed separately. For each variable, a repeated measures analysis of variance (ANOVA) was computed. The time of data collection was a within subjects factor, while group membership was a between-subjects factor. The respective information literacy performance indicator was used as dependent variable.

The performance on the knowledge test was analyzed first. The analysis revealed a significant main effect of the within-subjects factor (F[2,130] = 216.53, p < 0.01) and a significant interaction of the two factors (F[2,130] = 73.13, p < 0.01), what is depicted in figure 1. To analyze group differences, a t-test was computed for every data collection. At data collection 1, there was no difference between the groups (t[65] = 0.78, n.s.). At data collection two, a significant difference could be found (t[65] = 10.47, p < 0.01), indicating that group 1 outperformed the other group. There was no significant difference at data collection 3 (t[65] = 0.37, n.s.).

Next, the outcome scores of the search tasks were analyzed. The ANOVA revealed a significant main effect of the within-subjects factor F[2,130] = 45.77, p < 0.01) and a significant interaction of the two factors (F[2,130] = 5.45, p < 0.01). To investigate the pattern in more detail, t-tests were calculated to compare the two groups at each data collection. There were no significant differences at data collections 1 and 3 (t[65] = 1.33 and t[65] = 0.32, respectively). However, the two groups differed at data collection 2 (t[65] = 3.32, p < 0.01). Once again, group 1 outperformed group 2, as can be seen in figure 2.

Finally, the procedure scores of the search task were analyzed. The ANOVA revealed a significant main effect of the within-subjects factor F[2,130] = 148.46, p < 0.01) and a significant interaction of the two factors (F[2,130] = 37.38, p < 0.01). Once again, t-tests were applied to analyze group differences. There was no significant difference at data collection 1 (t[65] = 1.23, n.s.), but significant differences at data collections 2 (t[65] = 9.21, p <0.01) and 3 (t[65] = 3.21, p < 0.01) in such a way that group 1 scores higher than group 2, as is displayed in figure 3.

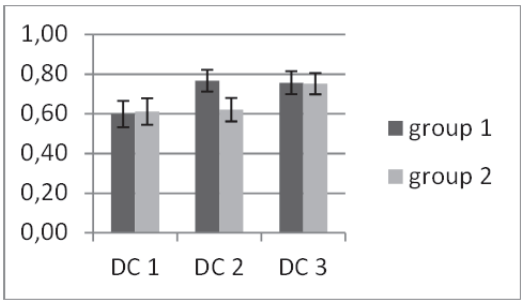


Figure 1: Mean scores (and standard deviations) on the information literacy test. DC = data collection.

Table 2: Mean scores (and standard deviations) for the outcome and procedure scores.

	data collection 1	data collection 2	data collection 3
outcome			
group 1 (*n*=37)	0.45 (0.19)	0.63 (0.19)	0.78 (0.17)
group 2 (*n*=30)	0.51 (0.16)	0.49 (0.13)	0.76 (0.18)
procedure			
group 1 (*n*=37)	0.43 (0.13)	0.81 (0.10)	0.78 (0.08)
group 2 (*n*=30)	0.47 (0.13)	0.54 (0.12)	0.70 (0.11)

Table 3: Mean scores (and standard deviations) for the information literacy knowledge test.

group	data collection 1	data collection 2	data collection 3
group 1	0.59 (0.06)	0.76 (0.05)	0.75 (0.05)
group 2	0.61 (0.06)	0.62 (0.05)	0.75 (0.05)

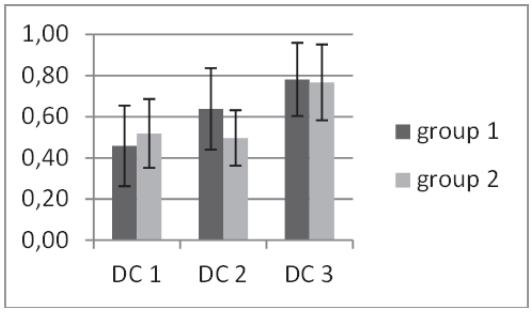


Figure 2: Outcome scores (and standard deviations) of the information search tasks.

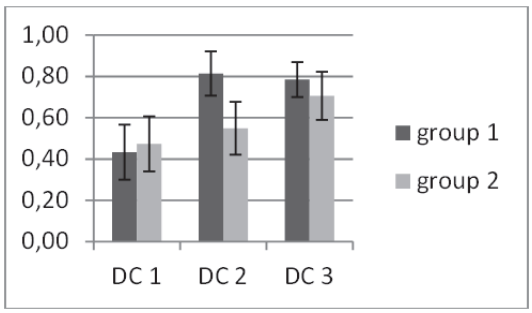


Figure 3: Procedure scores (and standards deviations) of the information search tasks.

3.4 Discussion

The results show that the hypotheses were confirmed. The first two hypotheses relate to the soundness of the evaluation instruments. First, the analysis of the search task difficulties indicated that the expected order of task difficulties could be verified empirically. This shows that those tasks requiring more competencies are also harder to solve for the participants, what can be seen as an indication of validity for the search tasks and the underlying task taxonomy. The second hypothesis, postulating that there would be significant correlations among all three information literacy performance indicators, could also be upheld. The fact that the correlations are far from perfect lead us to the conclusion, that all three instruments capture different facets of the concept information literacy. It is of significance that search task outcome and procedure scores both correlated significantly with the knowledge test, as this is a signal that the search tasks are valid, because the test had already been tested in a different study.

The third and most important hypothesis was that participation in the course improves information literacy. As can be seen in figures 1 to 3, participation in the course improved information literacy on all three performance indicators. To be exact, the hypothesis was that group 1 would outperform group 2 at data collection 2. This was hypothesized because at that time, group 1 had already taken the course, while group 2 had not. At data collection 3, there should not be any group differences left. This pattern could be observed when analyzing the scores on the knowledge test and the outcome scores of the search tasks. Analysis of the procedure scores displayed the same pattern at data collections 1 and 2. At data collection 3, however, group 1 still performed better, even though at this time both groups had taken the course. A more detailed analysis revealed that both groups improved; however, group 1 performance was enhanced stronger. No substantial explanation for this result can be found except that this might be random variations due to relatively small sample size. As mentioned above, group sizes were $n = 30$ and $n = 37$, respectively. Notwithstanding that participants had been assigned to one of the groups in a randomized way, it cannot be ruled out that some group 1 participants learned more due to individual differences. As this difference can only be observed on one of the three performance indicators, it can be ascribed to random variations. The hypothesis can still be confirmed as both groups had improved by taking the course. It should also be mentioned that there were significant differences between the groups at data collection 2. At this time, a

treatment group was compared to a waiting control group. The group differences show that participants did not become more information literate without training, corroborating the attribution of information literacy improvements to the participation in the course.

To sum up, the results show that the information literacy course is effective. So, this research adds to the field a blended learning course that is tailored to Psychology students and has been rigorously evaluated. In the future, it might be developed further by adding elements tailored to Master levels students. As this course was tailored to Bachelor students, the participants were mainly taught the essentials of seeking academic information. To make carrying out the course more resource efficient, it might also be an idea to leave out the classroom seminars. A further evaluation study might show whether this is equally effective.

REFERENCES

American Library Association. 1989. *Presidential committee on information literacy. Final Report* [online]. Chicago, IL: American Library Association. Available at: http://www.ala.org/acrl/publications/whitepapers/presidential [Accessed 30 September 2013].

Association of College and Research Libraries. 2010. *Psychology information literacy standards* [online]. Chicago, IL: American Library Association. Available at: http://www.ala.org/acrl/standards/psych_info_lit [Accessed 19 March 2013].

Carr, S. 2000. As distance education comes of age, the challenge is keeping the students. *Chronicle of Higher Education* 46(23), pp. A39-A41.

Center for Assessment and Research Studies. 2013. Information literacy test (ILT) [online]. Harrisonburg, VA: James Madison University. Available at: http://www.madisonassessment.com/uploads/Manual-ILT-2013.pdf [Accessed 29 January 2014].

Clardy, A. 2009. *Distant, on-line education: effects, principles and practices* [online]. Available at: http://files.eric.ed.gov/fulltext/ED506182.pdf [Accessed 29 January 2014].

Eisenberg, M. B. 2008. Information literacy: essential skills for the information age. *DESIDOC Journal Of Library & Information Technology* 28(2), pp. 39–47.

Julien, H. and Barker, S. 2009. How high-school students find and evaluate scientific information: a basis for information literacy skills development. *Library & Information Science Research* 31(1), pp. 12–17.

Kim, J. 2009. Describing and predicting information-seeking behavior on the Web. *Journal of the American Society for Information Science and Technology* 60(4), pp. 679–693.

López-Pérez, M., Pérez-López, M. and Rodríguez-Ariza, L. 2011. Blended learning in higher education: students' perceptions and their relation to outcomes. *Computers & Education* 56(3), pp. 818–826.

McClelland, D. C. 1973. Testing for competence rather than for "intelligence". *American psychologist* 28(1), pp. 1–14.

National Forum on Information Literacy. n. d. *What is information literacy?* [online]. Cambridge, MA: National Forum on Information Literacy. Available at: http://infolit.org/?page_id=3172 [Accessed 6 June 2012].

Noe, N. W. and Bishop, B. A. 2005. Assessing Auburn University Library's tiger information literacy tutorial (TILT). *Reference Services Review* 33(2), pp. 173–187.

Project SAILS. 2013. *Project SAILS information literacy assessment* [online]. Kent, OH: Kent State University. Available at: https://www.projectsails.org/ [Accessed 16 August 2013].

Shavelson, R. J. 2010. On the measurement of competency. *Empirical research in vocational education and training* 2(1), pp. 41–63.

Smith, J. K., Given, L. M., Julien, H., Ouellette, D. and DeLong, K. 2013. Information literacy proficiency: assessing the gap in high school students' readiness for undergraduate academic work. *Library & Information Science Research* 35(2), pp. 88–96.

Thaxton, L., Faccioli, M.B. and Mosby, A.P. 2004. Leveraging collaboration for information literacy in psychology. *Reference Services Review* 32(2), pp. 185–189.

Timmers, C. F. and Glas, C. A. W. 2010. Developing scales for information-seeking behaviour. *Journal of Documentation* 66(1), pp. 46–69.

Warwick, C., Rimmer, J., Blandford, A., Gow, J. and Buchanan, G. 2009. Cognitive economy and satisficing in information seeking: a longitudinal study of undergraduate information behavior. *Journal of the American Society for Information Science and Technology* 60(12), pp. 2402–2415.

SHORT PAPERS

A Study on the Permeation and Scope of ICT Intervention at the Indian Rural Primary School Level

Shrabastee Banerjee[1], Kalyan Sankar Mandal[2] and Priyadarshini Dey[3]
[1]*Department of Economics, The University of Warwick, Coventry CV4 7AL, U.K.*
[2]*Centre for Contemporary India Studies, Lund University, Paradisgatan 2, 221 00, Lund, Sweden*
[3]*Social Informatics Research Group, Indian Institute of Management Calcutta, Joka, Calcutta, 700104, India*
{shrabastee.b, kalyanmandal, priyadarshinidey27}@gmail.com

Keywords: ICTs in Education, Tribal Communities, Remote Teaching, e-Learning.

Abstract: The provision of education for all in India remains a distant dream, despite substantial amounts of government and state investment going into it. The objective of this study is to highlight an alternative learning model that makes use of the e-revolution that has proliferated into every aspect of our lives. Although there have been attempts to incorporate ICT into rural classrooms, most of the focus has been on video-based digitized learning and has not efficiently addressed the best ways in which learning can be achieved. Our aim is thus to design a model that not only makes e-learning effective, but replaces the under-qualified teachers in remote areas and allows for the free permeation of education in ways that might bridge the digital divide amongst students of various socio economic backgrounds. In this context our intervention focuses on a class of 16 students, 10 to 11 years of age (class 5) at Ma Sarada Shishu Tirtha, a school for tribal girls, located in Krishnanagar, West Bengal, India. The intervention involved a remote teacher delivering Math and English lessons in a class-room setting, (via the video conferencing software Skype, and PowerPoint Presentations) while also making the session interactive.

1 INTRODUCTION

Elementary education of underprivileged rural children residing in remote regions is the primary focus of the study, since only a robust elementary education system can solve chronic problems of low literacy and educational deprivation. Even though global demand tends to be focused on higher education and technical skills, there are backward linkages with elementary education in several ways. On the one hand, achieving universal elementary education is expected to raise productivity and incomes and strengthen the domestic market, seen as a condition for continued economic growth. On the other hand, the growing concern with basic education is seen as limited in the current economic scenario as it does not adequately consider the education and skill requirements needed to enhance productivity and incomes in a changing economy.

Development theory in recent years has taken note of the importance of education as an index of development of a nation, and with its myriad positive effects on the functioning of a society, the outreach of education to every stratum of society is a subject of great concern. The connection between universalisation of elementary education and human development is one of the most widely acknowledged issues in public discourse across the globe. Elementary education thus plays a pivotal role in reducing social inequalities by expanding human capabilities. (Sen, 1989)

Four features that have characterized India since Independence continue to characterize India's elementary education system: incomplete enrolment, inequalities, poor quality, and ineffective school performance. Further, despite aggregate improvements in education levels, glaring inequalities in basic education continue to persist. Disparities between regions (states) and across gender, caste, class, religious groups; and other marginalized sections of society continue to present the biggest challenge for policy makers and educationists (Dreze and Sen, 2002). For instance, about 53 % boys complete primary education compared to 34% girls. Recent interventions have helped in bridging the gender gap but the drop-out rate among girls, especially in primary classes, is still a cause for grave concern. This is reflected, for

instance, in the differential in the median years of schooling which is 5.5 years for boys compared to 1.8 years for girls.

Marginalized groups such as the scheduled castes (SCs) and scheduled tribes (STs) (two groups of historically-disadvantaged people recognised in the Constitution of India) as well as religious minorities continue to fall out of the loop of schooling. In addition, there are striking gender differentials in school attendance among children 6–14 years belonging to SC, ST, and rest of India. The gender differential is most pronounced among the ST communities—a gap of almost 12 percentage points (Kumar and Rustagi, 2010).

School performance is often marked by absenteeism, inadequately trained teachers with indifferent attitudes, non-availability of teaching materials, inadequate supervision, and little support. Many poor families, having lost faith in government schools, are forced to send their children to private schools even when they have access to 'free' public schools. Several cases of discrimination are reported – against girls, against children belonging to socially disadvantaged and minority communities, and against the poor. Corporal punishment is common and many children are afraid of going to school for fear of being beaten, if not publicly ridiculed or rebuked by teachers and other students. (Kumar and Rustagi, 2010)

Although the proportion of GNP allocated to education (revenue and capital accounts together) has grown from a very low level of 0.67 % in 1951-52 to reach the all-time high of 4.4 % in 2001, mass illiteracy continues to flourish. Education is by far the largest capitalized space in India with $30bn of government expenditure (3.7% of GDP; at global average), and a large network of around one million schools and 18,000 higher education institutes. Yet, the public education system is 'insufficient' and 'inefficient'. A break-up of government expenditure shows that only a miniscule 0.82% component goes towards capital expenditure. A whopping 80% of the revenue expenditure on teachers' salaries leaves little to be spent on infrastructure creation, which eventually translates into 'ineffective' infrastructure/ quality of education. (Vora and Dewan, 2009). It indicates that a higher allocation of funds on education is not the only criterion to solve our literacy problems.

Although, various schemes and programmes have been started by the State Governments and the Ministry of Tribal Affairs to promote universal primary education among tribals such as- scheme for establishment of Ashram schools in tribal sub-plan areas, scheme for establishment of boys hostels for the Scheduled Tribes, scheme for construction of girls hostels for the Scheduled Tribes, and scheme for development of Primitive Tribal Groups (PTG), but in reality very few of them have percolated down to the tribal and benefited them. Many of the programmes did not benefit the tribal community because the programmes were not contextualized and localized considering regional, geographical and physical differences and barriers. (Kumar, 2008)

ICTs are a potentially powerful tool for extending educational opportunities, both formal and non-formal, to previously underserved constituencies—scattered and rural populations, groups traditionally excluded from education due to cultural or social reasons such as ethnic minorities, girls and women, persons with disabilities, etc.(Roy, 2012)

ICT based distance education in India has primarily been confined to university-level education and is often considered being sub-par when compared to traditional courses. The long-term purpose of the present study is to develop a sustainable model of distance-learning that is cost-effective, and leads to a more fulfilling learning experience for the children.

The objective of this study is to highlight the importance of working with constrained resources, and proposes an alternative learning model that makes use of the e-revolution that has proliferated into every aspect of our lives. The study examines whether it can be utilised to overcome infrastructural bottlenecks in the provision of quality education in inaccessible rural areas, and thus bring about a concomitant increase in the population's educational levels. The internet-based social media revolution of the present age has the possibility of transforming education and knowledge as one knows it. By exploiting this, it may be possible, via a model that combines the correct elements of distance learning, to bring forth a knowledge revolution and spread education to the remotest corners of the country. Our aim is thus to design a model that not only makes e-learning effective, but replaces the under-qualified teachers in remote areas and allows for the free permeation of education in ways that might bridge the digital divide amongst students of various socio economic backgrounds.

2 ICT BASED LEARNING

2.1 Computer-Assisted Learning

The wide use of Internet has affected the methods of education, giving us the concept of global classroom where any number of students can interact with each other at any time. ICT is currently widespread in schools and colleges of developed countries (DG Communications Networks Report, 2013). In the developing countries, it is being primarily used to support education in the form of "E-Learning" applications. There is still some amount of scepticism among the schools in adopting ICT because it is still a relatively new feature and its impact on student achievement is not known. Also such technological solutions are thought to be not very cost effective (Rajesh, 2003).

Although there are several studies that indicate how ICT can benefit various stakeholders, what is certain is that ICT is not the ultimate solution rather an aid to achieve the maximum that can be derived out of educational experiences at school. While the debate on the impact of ICT on learning outcomes still goes on, there seems to be a consensus that ICTs in education can increase access to information inside a classroom as well as increase motivation and efficiency throughout the system. (Haddad and Jurich, 2002) (InfoDev Report, 2010)

Government of India, as part of its 11th Five Year Plan, continues to support federally sponsored scheme, known as "Sarva Shiksha Abhiyan, (SSA)" with the objectives of providing school education to every child between the age of 6 and 14 years and improving the quality of school education in the country. Under this policy GOI aims at fostering ICT based Education in India. Under the Sarva Shiksha Abhiyan, several Computer Aided Learning (CAL) programmes have been created by developing multimedia content.[1]

The objectives of CAL are as follows:

1. To facilitate effective delivery of curriculum content
2. To act as an effective supplement for teachers to improve learning levels in the school since it facilitates practical and experimental learning.
3. To serve as a means to attract children to schools with the multimedia i.e. audio-visual form of learning on various subjects of classroom teaching and thus hold their attention, hence tackling the challenge of dropouts and achievement of enrolment.

[1] (www.ssa.tn.nic.in/Docu/CAL%202011-ss.pdf)

SSA has introduced many CAL projects across states in India. 27,289 schools with about 5.3 million students were beneficiaries of this program.[2] The two major ICT related projects in schools are Multimedia in K-12 (for *affluent* private schools) and ICT in public schools. The leading players are Educomp Solutions, Everonn and NIIT. Educomp which is the pioneer in this field has a product called "Smart Class". Smart class is a product offered under Multimedia in K-12. The objectives of this are as follows:

1. To use multimedia modules to effectively teach students across Kindergarten to 12th grades.
2. Content delivery through multimedia i.e. digital content and LCD TVs/Projector assisted whiteboards. It has been universally accepted that this content delivery scheme improves teaching and learning experience.

However, the focus of these efforts was more towards computer-aided learning than Distance e-Learning . Distance E-Learning is the combination of Distance Education and e-Learning which is characterized by the extensive use of Internet enabled web technology in the delivery of education and instruction and the use of synchronous and asynchronous online communication in an interactive learning environment or virtual communities, in lieu of a physical classroom, to bridge the gap in temporal or spatial constraints (Garrison, 2011). A recent development in distance e-learning is Massive open online courses (MOOCs), aimed at large-scale interactive participation and open access via the web or other network technologies.

But, MOOC is usually non-interactive, i.e. the students will be viewing the recorded videos by the instructors, followed by off-line question-answering. Our focus in this study is to use Distance e-Learning in a synchronous mode, which implies real-time interaction with fixed-time-fixed-place virtual classroom over the Internet for primary students.

The McKinsey report (Madgavkar et al, 2012) states that there will be 330 million Indian users of the internet in 2015, thus making it the second largest connected population in the world. Rural access to education can be vastly improved by means of exploiting this revolution, and creating a networked virtual classroom system.

2.2 ICT Interventions in Rural Classroom

Some efforts in the sphere of remote learning

[2] mhrd.gov.in/ict_school

already undertaken in rural India are summarised as follows:

Hole-in-the-Wall (HitW): Hole-in-the-Wall is a 'Shared Blackboard' which children in underprivileged communities can collectively own and access," to express them, to learn, and to explore together." In an experiment conducted first in 1999, a computer was placed in a kiosk created within a wall in a slum, and children were allowed to use it freely. The experiment aimed at proving that children could be taught by computers very easily without any formal training. Sugata Mitra, the inventor, termed this as Minimally Invasive Education (MIE). The experiment has since been repeated in many places, HIW has more than 23 kiosks in rural India. This work demonstrated that groups of children, irrespectively of who or where they are, can learn to use computers and the Internet on their own with public computers in open spaces such as roads and playgrounds, even without knowing English.

Pratham's Computer Aided Learning: This is a school based program which caters to school going children from 6-18 age groups with about 40% children in secondary school age. The objective of this program is: 1) To impact children's basic learning levels using IT and 2) to give them relevant IT knowledge and skills. Through its school based computer labs, this program reaches out to close to 90,000 children across 7 states. This program also tries to improve schools performance by encouraging them to adopt various IT solutions like MIS, Database etc and to get teachers to adopt technology through teacher training. The curriculum includes software developed by Pratham in local languages that helps build the reading and math skills of the children.

Some other interventions have been observed in the sphere of remote teaching and learning, though the focus of the following is not rural:

1. Board of Open Learning School: This was intended to be a medium for making distance learning easy for students who are unable to attend regular classes.[3]
2. iPerform: It is a web-based e-learning platform that works as an extension of the classroom and provides a personalised self-learning environment for students. It has been designed to meet differential learning styles, engaging students in activities like homework, revision and even help in preparation of exams.[4]

However, all these methods are mostly focusing on asynchronous learning based on preloaded course material and scope for real-time interaction are absent. As discussed, our focus is synchronous, i.e. virtual classroom over internet, where teacher should be present at remote location during teaching..

[3] www.bolsd.in

[4] www.iperform.classteacher.com

2.3 Critical Evaluation of ICT Intervention

A variety of constraints (Rajesh, 2003) dodge India's efforts to deploy technology in education:

1. Policy and Government commitment are existent but all is lost on the road to implementation. Educational projects, set up by conventional governments as part of a broad educational agenda, tend to reflect the conventionalism of existing institutions with their hierarchical and bureaucratic systems of administration when the need is for creative and innovative management.
2. Access and availability of technology is another issue as various implementing agencies that need to cooperate are not coordinating enough. Lack of stable electric power, non-existent or unreliable telecommunication lines and a mismatch between funding allocation and actual needs all add to the problems.
3. Sustainability is also a major obstacle, with many initiatives failing because Governments have not anticipated the cost of maintenance and upgrading of technology and services.
4. Geographic, Demographic and Climatic factors affect the access, reach and implementation of infrastructure
5. Ethnic and Cultural factors have an influence on the teaching or pedagogical methods. Language barriers are major obstacles as well.

Most importantly, however, the proper means for implementing an ICT-backed educational model has not been developed, with undue attention being paid simply on the process of setting up ICT labs. Our study tries to approach this issue in a manner that can solve these basic problems to bring about the complete benefits of an ICT intervention.

3 THE INTERVENTION AT A TRIBAL SCHOOL IN INDIA

The long-term purpose of the study would be to develop a sustainable model of distance-learning that is cost-effective, and leads to a more fulfilling learning experience for the children. The study focuses on a class of 16 students of class V and 28

students of Class I of Ma Sarada Shishu Tirtha, hailing from the tribal community in Krishnanagar, India. The undertaken intervention involved a remote teacher delivering Math and English lessons, (via the video conferencing software Skype, and PowerPoint Presentations shown during the video conference) while also making the session interactive and addressing individuals students directly by name, which brought the additional dimension of familiarity and added to the success of the intervention. The survey finds sufficient evidence to suggest that ICT intervention can significantly enhance and improve their academic attainment and leads to a more fulfilling learning experience for the children. Further, the present model is not one of distance learning in the conventional sense, in that it combines elements of traditional classroom teaching in a virtual form, enhancing it with tools as and when required. The main aim is to design a model that not only makes e-learning effective, but to replace the traditional model to save recurring expenses on local teachers and allow for the free permeation of education in ways that might bridge the digital divide amongst students of various socio economic and other geographical barriers.

For a preliminary understanding of the current level of the children, tools developed by Pratham, an NGO based in Mumbai, India were used. The ASER Centre of Pratham seeks to use simple yet rigorous methods to generate evidence on scale on the outcomes of social sector programs. (Vagh, 2009). We have also conducted written and oral examination to assess the effect of ITC intervention.

The before study comprised of usage of the ASER tools, a questionnaire to assess the academic achievement of the children. The children's past school results were also used to assess the baseline standard. Upon using the ASER tools, it was discovered that, while the children were fluent readers in Bengali, their mathematical ability was below average, with only 50% of the students being able to solve Class 2 level division problems. The questionnaire further revealed that English and Math remained a major concern for the students.

During the intervention, Math and English lessons were conducted for class V and English lessons were conducted for class I remotely, via the video conferencing software Skype, and PowerPoint Presentations shown during the video conference. (The intervention was preceded by a short introduction session to make the students feel at ease.) After the intervention, a follow-up questionnaire was used using some exercises covering the topics taught via ICT, to gauge their interest, retention and possible improvement.



Figure 1: Children Being Taught Remotely (Inset: The Remote Instructor).

3.1 Methodology

The research is Quasi-experimental in nature and has undertaken before after studies to evaluate the impact of the intervention. The before study entailed the students being ranked on the basis of their performance in class (more specifically on the basis of school examination results). The ASER survey questionnaire for wellbeing assessment and educational assessment of English has been used. Next, the students have been ranked on the basis of their performance after the intervention in English through ICT. The impact analysis comprises of assessments and interactive observation with the students.

3.2 Evaluation

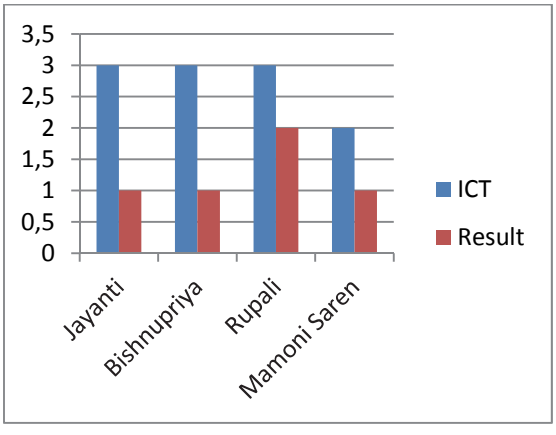


Figure 2: Relative performance of four students of Class-V based on Results of school exam. and evaluation done after ICT intervention

It can be seen that those who had been near the bottom of the class benefit the most from this intervention (fig.2). A majority of the others perform at the same level as their class performance. Among the five students of class V who have a normalized score of 1 in their school exams, three (i.e. 60%) achieve a normalized score of 3 in the evaluation conducted post-ICT intervention. Although the sample size is small, this is a significant finding.

Fig 3. shows the examination performance of 28 Class-I students in (i) the regular school examination before remote teaching, and (ii) an examination taken after a month-long remote teaching intervention in English (20 hours total). The students have been indexed on the basis of their rank in the school exam. The chart shows marked improvements in student performances. Specially, the students who received very poor marks in school examination (student number 24 to 28) have shown significant improvement in their English language skills. The classes were conducted using Google plus and the teaching learning materials were video based and PowerPoints. The method of teaching was interactive and it was made sure that the children interacted with the instructor.

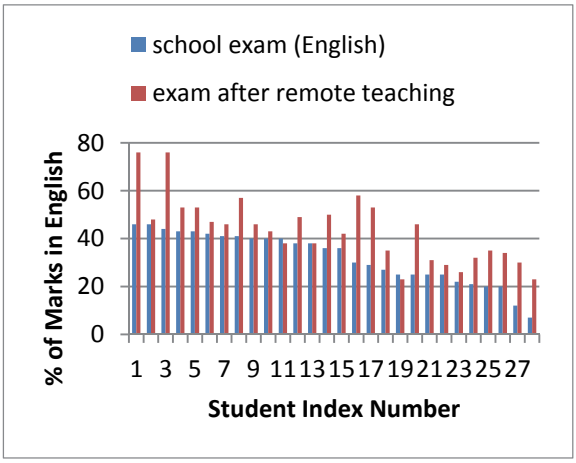


Figure 3: Performance Improvement of Class-I Students after ICT-enabled Remote English Teaching.

The intervention has generated positive results both in nature of change in attitude towards technology and learning outcomes. The children were aged between 6 to 11 years and it was earlier assumed that it would be difficult for them to cope with remote teaching due to their lack of attention span. However during the intervention it was observed that the children on the contrary did not lose their attention. The social difference that if often created between teacher and student due to several disparities was not observed in this context, the children were quite aware of the instructor's presence online but they did not have any hesitation in interaction. There was no fear of punishment although the instructor observed discipline in the class, but the children did not shy away from the instructor. Interactive sessions created a positive learning environment for them, leading to increased interest in the topics taught.

It can be deduced that technological aids, along to low/no teacher absenteeism and low social difference has lead to creation of a better learning environment. The teacher student ratio being comparatively low, and the regular interactions with the teacher also lead to faster learning of concepts. Video based teaching created better cognitive understanding of the topics taught. Children were able to freely interact amongst themselves, and this also created a positive interest in the topics taught. Lastly the expectancy created in the children, increased interest for both instructor and the children in the classroom.

To summarize, one can make the following hypotheses about this intervention and, based on the data obtained, examine whether the nature of results for each has been positive or negative:

1. **The Children Enjoy the ICT-enabled Classes More in Comparison with Their Usual Class:** We have received **positive response** from most of the students on the basis of the questionnaire and observations. 70% commented that they found the pictures/videos shown to be very engaging and interesting. This is natural, since an audio-visual presentation is more engaging for children and keeps children attentive longer than a routine class.
2. **An ICT-based Remote Teaching System Serves to Bring about Higher Teacher Acceptability:** We have received **positive response** from most of the students on the basis of the questionnaire and observations. 89% students stated that they found the remote teacher to be highly acceptable. Further, since there is no physical threat from a "remote" teacher, the children feel more relaxed in presence of the teacher. 81 % students have stated they were mostly able to understand the remote teaching. There were no negative responses. Hence, acceptability is generally high during remote class.
3. **An ICT-based Remote Teaching System Serves to Bring about an Improved Performance in Many, in Comparison to School Performance:** Although the data set is

small, and time of intervention was short to draw any definitive conclusion about this, fig. 2 and fig.3 nonetheless shows marked improvements in student performances. Specially, the students who received very poor marks in school examination have shown significant improvement in their English language skills This could be because the ICT-aided material that was taught was better understood and retained by the children.

4. **The Children Will Fear or Be Skeptical of New Technology:** We have received **negative response** from most of the students on the basis of the questionnaire and observations. Rather, the children have shown a reasonably ready acceptance of it, and 81% interviewed have stated that they are open to change and would like to be taught in a different way.
5. **The Children Would Offer Some Resistance in Opening up to Outsiders:** We have received **negative response** from most of the students on this. Being a close-knit tribal community at a remote village, with little interaction with the outer world, it was assumed that the children would not be open to the idea of complete strangers instructing or interacting with them. However, not only did the children show incredible openness to the idea, they also engaged very freely in all interactions and showed extreme enthusiasm.

4 DISCUSSION AND CONCLUSION

Teachers are expected to enrich a child's learning and schooling experience. But this is often not the case in rural India. Studies have shown that teachers frequently beat children, terrorize them, and humiliate them publicly. Many forms of discrimination and biases enter the classroom. A recent survey of rural schools, for example, carried out in West Bengal found disturbing evidence of primary school teachers showing much less regard for the interests of children belonging to Scheduled Castes. Teachers tended to perceive themselves as belonging to a different and higher class, often the result of earning much higher incomes than most parents. They rebuked children for not coming properly dressed to school, for being obviously dirty, for being stupid because they belonged to a certain community. Children were ridiculed for their eating habits. In some instances, they were made to sit separately. (Kumar, Rustagi, 2010)

Better-off sections belonging to higher castes are able to send their children to the fee-charging private school, which they can afford, for better quality of education. The poor belonging to lower castes, not being able to afford private school, are destined to send their children to inferior quality government schools. Studies indicate that first; most of the children enrolled in private schools are from general caste group, whereas most of the SC and ST children attend government schools (Aggarwal 2000; Kumar et al 2005; Mehta 2005; PROBE 1999).

ICT is a significant step in this direction, which eliminates barriers of regional and gender-based discrimination by facilitating a high degree of permeability of education. Since this was a pilot study, the scope was limited by several constraints, but we certainly plan on using more sophisticated tools of e-learning for future interventions.

With regard to the survey conducted at Ma Sarada Shishu Tirtha, one probable reason for the ready acceptability of ICT, and the improvement in performance, could be that, if the teachers in a face to face teaching situation are from non-tribal background, a 'social distance' may be created which operates as a hindrance in learning. This is however more applicable with respect to weak students. ICT teaching being neutral to 'social distance' is more helpful for disadvantaged groups. It is clear that a long-term intervention is in order, and could prove to be immensely beneficial if implemented in a systematic manner.

While "digital inclusion" programs usually tend to focus on teaching people the usage of computers and the internet, the approach presented in this study takes a different perspective. According to this approach, the entire process of learning should be oriented towards developing abilities to connect to the global knowledge network of the cyber-world with a specific social context and purpose in mind. This makes it possible to improve the underdeveloped communities by obtaining social inclusion through digital inclusion, and creating self-sustainable forms of social development.

REFERENCES

DG Communications Networks & Technology (2013). Survey of Schools: ICT in Education-Benchmarking Access, Use and Attitudes to Technology in Europe's schools (A study prepared for the European Commission)

Dreze, J. (2003). Patterns of Literacy and their Social

Context, in Das V. et al (ed). *The Oxford India Companion of Sociology and Social Anthropology, 2*, Oxford University Press, New Delhi.

Dreze, Jean and Amartya Sen (2002). *India: Development and Participation*, Oxford University Press, New Delhi.

Garrison, D.R. (2011, 20 May). *E-Learning in the 21st Century: A Framework for Research and Practice*. New York: Taylor & Francis. ISBN 0-203-83876-9

George, K.K. and Reiji Raman (2002). Developments in Higher Education in India-A Critique, working paper No.7, Centre for Socio-economic & Environmental Studies.

Gnanasambandam, Chandra, Anu Madgavkar, Noshir Kaka, James Manyika, Michael Chui, Jacques Bughin, Malcolm Gomes (2012). Online and upcoming: The Internet's impact on India, McKinsey.

Garg Manisha, Mandal, Kalyan Shankar, (2013). *Mid-Day Meal for the Poor, Privatised Education for the Non-Poor*, Economic & Political Weekly, Vol XLVIII no 30.

Kumar, Anant, (2008). *Education of Tribal Children in Jharkhand: A Situational Analysis,* Jharkhand Journal of Development and Management Studies, XISS, Ranchi, Vol. 6, No.4 (XXV).

Haddad, Wadi and Sonia Jurich (2002). ICT for Education: Potential and Potency, in *Technologies for Education: Potentials, Parameters, and Prospects*, eds. Wadi Haddad and A. Drexler (Washington, D.C.: Academy for Educational Development), 28-40.

InfoDev Essay II: for PriceWaterHouse Coopers (2010). ICT in School Education (Primary and Secondary)

Kumar, A. K. Shiva and Preeti Rustagi (2010). Elementary Education in India: Progress, Setbacks, and Challenges, Oxfam India working papers series, OIWPS – III

Rajesh, M (2003). *A Study of the problems associated with ICT adaptability in Developing Countries in the context of Distance Education,* Turkish Online Journal of Distance Education-TOJDE.

Roy, Niraj Kumar (2012). *ICT –Enabled Rural Education in India*, International Journal of Information and Education Technology, Vol. 2, No. 5.

Sen, Amartya and Jean Dreze (2013). *An Uncertain Glory: India and its Contradictions.* Allen Lane.

Sen. Amartya, (2003). Development as Capability Expansion. In: Fukuda-Parr S, et al Readings in Human Development. New Delhi and New York: Oxford University Press; 2003.

Vagh, Shaher Banu (2009). Validating the ASER testing tools: Comparisons with reading fluency measures and the Read India measures. Technical Paper, ASER Centre, New Delhi 110 029, India.

Vora, Nikhil and Shweta Dewan (2009). Indian Education Sector, IDFC SSKI India Research.

Board of Open Learning School: www.bolsd.in/Computer Aided Learning under SSA: www.ssa.tn.nic.in/ Docu/CAL%202011-ss.pdf

ICT in schools: http://mhrd.gov.in/ict_school

iPerform: www.iperform.classteacher.com

"Smoking Does Not Make You Happy"
Unlearning Smoking Habits Through Mobile Applications on Android OS

Răzvan Rughiniş[1], Ştefania Matei[2] and Cosima Rughiniş[2]
[1]*Department of Computer Science, University Politehnica of Bucharest, Bucharest, Romania*
[2]*Department of Sociology, University of Bucharest, Bucharest, Romania*
razvan.rughinis@cs.pub.ro, {stefania.matei, cosima.rughinis}@sas.unibuc.ro

Keywords: Smoking Cessation, Android Applications, Quitting Advice, User Profile.

Abstract: We analyze in-depth five smoking cessation apps on Android OS, examining how they teach users to quit smoking and what they learn from users. Apps advise would-be ex-smokers how to perceive the world, how to deal with their emotions, and how to act on their bodies and environment. Still, they learn little from their users, and even less from the scientific literature on smoking cessation. We discuss the potential for improved customization of advice to users' profiles and we propose a simple inventory of online scientific resources as a starting point for developers looking to create better apps.

1 INTRODUCTION

There has been a recent increase in the use of technologies for smoking cessation interventions. Online programs and sites assist people who want to quit smoking (Quit Net, Smoke Free, Make Smoking History, Verizon Health Zone, Dit Digitale Stopprogram etc.). SMS Services of support and counseling (Guerrilla Interactive SMS Service, MiQuit, Mobily Naqa, etc.) send messages tailored to the situation of the recipient, giving advice and encouragement.

A current advancement consists in mobile applications designed to help people unlearn their smoking habits, and learn a new identity with its associated routines: the ex-smoker. They are available on multiple platforms and mobile operating systems (iPhone, Android, Windows 8 etc.), being released by a variety of actors, such as independent developers, software development companies, non-profit organizations for public health, or public institutions.

Smoking cessation applications are, essentially, attempts at 1) unlearning through re-interpretation of smoking, and 2) assembling a new set of daily life routines to support the transition to a new identity. Their first challenge is to encourage users to reframe smoking from a positively charged behavior (understood in terms of prestige in group, belonging to a group or social class, sign of capital assets, personal and psychological relief etc.) to a damaging action (as regards risks for self and others of premature death, disability, low fitness and degraded beauty; financial costs; low environmental comfort etc.). Their second challenge is to support users through withdrawal symptoms and craving moments, thus avoiding a relapse.

This paper reports a case study consisting of in-depth, qualitative analysis of five smoking cessation applications on Android OS. We focus on communication with users, addressing the following questions: What do apps teach users? What do apps learn from users? What evidence-based resources are available for developers to improve the fit between users' profiles and app messages?

2 SMOKING CESSATION APPS

The market of smoking-related applications contains both pro-smoking software (promoting tobacco products and smoking) and anti-smoking applications (framing tobacco consumption as negative). Even if there are some smoking-supporting applications (e.g. smoking simulators) which claim to help people quit, and could possibly be of assistance in some instances, most of the pro-smoking application are designed as brand advertisements, as wallpapers or widgets, and as claims to protect the right of smoking enthusiasts (BinDihm et al., 2012).

A comprehensive review of smoking cessation

applications for mobile phones show that most quitting applications do not adhere to key medical guidelines: "Few, if any, apps recommended or linked the user to proven treatments such as pharmacotherapy, counseling, and/or a quit line" (Abroms et al., 2011). The applications propose what we can term a "lone rider" model of the user: a heroic vision of quitting unassisted by medicine or professionals. A similar conclusion is advanced by a review of alcohol quitting apps (Cohn et al., 2011).

We start from this observation and we examine in-depth advice and tips offered by five smoking cessation apps which are, in our evaluation, typical for apps that are based on a combination of quantification, gamification, community, and textual communication, the most frequent type on the Google Play market. We then discuss our findings by pointing out missed opportunities for personalization of advice, and missed opportunities for learning from the available online scientific discussion on smoking cessation.

3 DATA AND METHOD

For our case-study we selected for in-depth analysis five smoking cessation applications on Android: [1] *Stop Smoking* (developed by Konrad Müller, last updated: November 2011), [2] *Quit Smoking Coach* (developed by Studios BrainLag, last updated: August 2013), [3] *Kwit - Quit Smoking is a Game* (developed by Geoffrey Kretz and Nicolas Lett, last updated: August 2013), [4] *Stop Smoking Assistant* ("As If" Licence, last updated: September 2012) and [5] *Exsmokers iCoach* (developed by Saatchi & Saatchi Brussels for the European Commission, April 2013).

Through this selection we cover *application variability in terms of design features* (Abroms et al., 2011): the selected apps illustrate quantified communication, textual advice and tips, games and gamification, and community support. We have also selected for analysis apps among those that include *richer textual messages for users*, so that we could examine their patterns of communication. We have identified and analyzed a total of 498 pieces of advice and panic tips provided by these five apps. We excluded tools that are based on hypnosis or subliminal messages because depend on strong assumptions, addressing a specific niche of users.

Based on our previous exploratory survey of anti-smoking apps, we consider that the five selected solutions are broadly representative of the available smoking cessation apps found on Google Play, integrating the most frequent options available in this type of work. Our case study serves to illustrate significant patterns and to open a discussion on the potential for improving advice for people who attempt to quit.

4 RESULTS

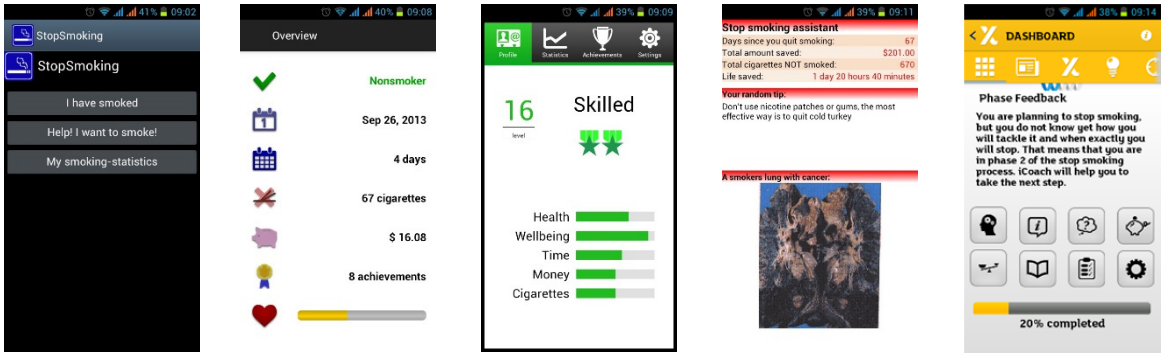
Smoking cessation applications often propose one or several frames for the communication with their users. Some of the most common frames are:

1) Coach or assistant: the application guides the user through the difficult venture of quitting;
2) Game: the application offers a gameful environment in which smoking cessation is a quest, or in which ex-smokers can enjoy games as rewards for their abstinence;
3) Community: the application facilitates an online support community of ex-smokers.

Communication between the application and users mostly occurs in the coach/assistant frame. It is bi-directional, although the two flows are, as a rule, clearly separated. Users are required to provide for the application a profile of their smoking habits, which are their most important identifier: this information typically includes the number of cigarettes smoked per day, time spent while smoking a cigarette, the number of cigarettes per pack, and the cost of a pack. Additional information on smoking inception and type of cigarettes may also be solicited. Users are also expected to honestly communicate each cigarette smoked; if the user has identified herself / himself as an ex-smoker, reporting a relapse usually resets all indicators and progress measures to zero. Last but not least, users are expected to ask for tips in case of cravings and feelings of panic. In turn, the application provides two different types of information:

1) A quantitative representation of the user's transformation and foreseeable future, with a focus on health benefits and financial savings (see app front pages illustrated in Figure 1); this representation usually include a combination of numbers and percentages, progress bars and other charts;
2) Advice, consisting of motivational quotes or panic tips; some applications also include motivational graphic images illustrating smoking-induced damage.

Therefore, apps learn from users, and teach users how to deal with smoking cessation. As we argue below, smoking cessation apps do a minimal work in learning from users, as user information is only



[1] Stop Smoking [2] Quit Smoking Coach [3] Kwit – Quit Smoking is a Game [4] Stop Smoking Assistant [5] Exsmokers iCoach

Figure 1: The front page of smoking cessation applications.

employed in the quantitative engines that generate numerical indicators and predictions, with little relevance for advice and tips.

4.1 Teaching Users

Apps are in charge of assisting users in a transition from a life structured by smoking routines, to a smoking-free life, through a period of withdrawal symptoms and craving. To this purpose, applications teach users to perceive the world differently, to manage their emotions and see themselves in a new way, and to engage in specific behaviors to minimize risks of relapse. We examine below in detail how apps teach users to deal with the experience of smoking cessation on four domains: perceptions, emotions, bodily actions, and environmental contexts (Figure 2).

As regards *perceptions*, there are four teaching areas:

1) Perceiving the future: users are presented with information about the distant future, especially concerning their improved health status in 10 to 20 years; users are also faced with descriptions of the closer future – a period free of cravings and dependency, in which the benefits of cessation are already visible and the costs have become memories;
2) Perceiving one's body in time: users are prompted to consider the gradual transformations of their bodies, while and after giving up smoking;
3) Changing one's perception of cigarettes as objects of consumption.
4) Managing one's perception of time, through time work (Flaherty 2003);

As regards one's emotions and identity, apps purport to teach users about:

1) Managing one's emotions, in a process of *emotion work* (Hochschild 1979);
2) Changing one's definition of oneself, in a process of *identity work* (Schwalbe & Mason-Schrock 1996).

Last but not least, apps also teach users to act on their *body*, shaping physical needs and sensations, and to be aware of one's context and act on the material and social *environment*.

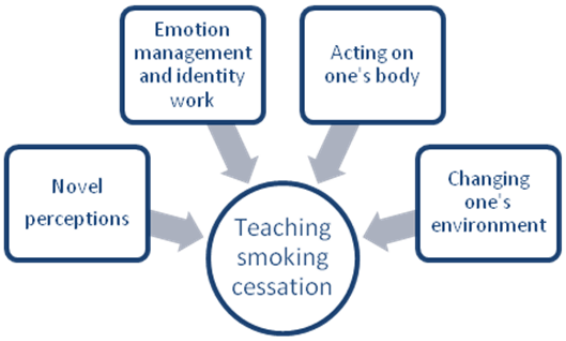


Figure 2: Four domains of app advice for users.

4.1.1 Perceiving the Future

The quantitative representation of users' progress is the most important tool for visualizing the future. Applications take over the role of a 'magnifying glass', compensating for users' myopia as regards future risks and benefits. The future is materialized as: the right end of a progress bar that fills gradually, the prospect of full recovery from smoking damage after 15 years of non-smoking, life-years gained through non-smoking, and through various metrics that display improved health and life expectancy.

Applications attempt to direct users to mind the future not only through indicators, progress bars and charts, but also through advice and panic tips (see Figure 3; in all Figures presenting application advice, the source of each quote is indicated in

square brackets, referring the application list from the Data and Methods section). Users are put in the situation to foresee their future health, but also their change of mind and emotions.

> Ten years after quitting your lung cancer risk will be reduced to 50%. [A1]
> In a few months, you will no longer think about smoking. [A1]
> When you haven't smoked for a month or more you will realize how stupid it was to spend all that money on an addiction that was literally killing you! NEVER AGAIN! [A4]
> Quitting keeps you young! Research has proven that smoking affects your body's cells, accelerating the aging process.[A5]
> The most difficult thing is to resist the first few weeks, and mainly the first days. It will be easier as time passes. [A3]

Figure 3: Teaching users: Perceiving the future.

4.1.2 Perceiving One's Body in Time

Users are instructed to attend to gradual improvements in their bodily well-being. This learning project is one of the most difficult to sustain, since it is subject to the possibility of direct disconfirmation: smokers are expected to experience all these improvements but, occasionally, it is possible that they do not occur as planned. Applications address this vulnerability, at least in part, by not soliciting users to actually observe their experiences: instead, the transformations are formulated as known-facts, in the present tense or future tense. These statements actually instruct the user how to perceive her bodily functions. Health improvements cannot be possibly observed on an hourly basis – but they can be reflected upon as such. It is also likely that respiration will not improve perceivably in three days – but the improvement that occurs is rendered visible through the application.

> Without smoking your health will increase from hour to hour. [A1]
> Three days after last cigarette the function of your respiratory system is improving. [A1]
> 48 hours after the last cigarette smells and tastes are being experienced more intensively. [A1]
> Your teeth, your breath and your skin thank you. Smoking leads to tooth loss, gives you bad breath and a shallow complexion. [A3]
> Think of your sex life! Smoking influences blood circulation in the genitals, which can lead to diminished ability in bed. [A5]

Figure 4: Teaching users: Perceiving one's body.

While in the previous section we have examined the applications 'telescopic' function for seeing the future, that compensates for users' time discounting (Frederick et al. 2003), this area of advice positions the application as a 'microscope', useful for revealing users the minute, gradual, real-but-unobservable changes in their bodies.

4.1.3 Changing Perceptions of Cigarettes

Applications also attempt to change users' perception of cigarettes themselves as objects of consumption. These messages are comparatively rare: we have identified only 4 out of a total of 498 (see Figure 5).

> Each cigarette shortens your life expectancy for 8-10 minutes. [A1]
> Light cigarettes will not help… you will simply smoke more of them or inhale deeper to get same effect. Just don't smoke. [A4]
> Think of the many poisons in a cigarette: ammonia, acetone, methanol,... still want to light up? [A5]
> By panicking, you give the cigarette too much power. Some perspective: a cigarette is nothing more than some tobacco wrapped in paper. Nothing special. [A5]

Figure 5: Teaching users: Perceiving cigarettes.

4.1.4 Changing Perception of Time

Time work refers to people's actions undertaken to transform the experience of time – that is, either to speed it up or slow it down, or modify its rhythms and flavors (Flaherty 2003). This is particularly relevant for ex-smokers: applications provide instructions for managing the objectively short but subjectively long periods of craving (see Figure 6).

> Delay: Craving usually passes quickly. Tell yourself that you must wait "just 10 more minutes" until craving passes. [A2]
> The urge to smoke is due to the lack of nicotine, it doesn't last more than 5 minutes. Stand firm and drink a big glass of water. [A3]

Figure 6: Teaching users: Time work.

Users are made aware of the temporary and brief nature of cravings, and they are taught a variety of actions that could serve to distract attention and make time pass more quickly.

4.1.5 Emotion Work

Hochschild (1979) introduces the concept of 'emotion work' to refer to the actions through which

people conform to emotion rules, adjusting their emotional displays and experiences in accordance with the situation. Emotion work is a central element of advice for smoking cessation applications, as recent ex-smokers are expected to confront many negative feelings.

The applications use several generic resources: a focus on the pride of transformation, positive emotions induced through gratifications, acceptance of negative emotions, and plain reframing: "smoking does not make you happy. It is simply very addictive [A1]" (see Figure 7).

Actually, smoking is no longer fun for you? [A1]
Smoking does not make you happy. It is simply very addictive [A1]
You will be very proud of yourself as an ex-smoker [A1]
Buy something nice: It's ok, you are saving a lot of money though not smoking. Reward yourself. [A2]
Think positive thoughts about how awesome it is that you are quitting smoking and getting healthy. Be patient with yourself. [A3]
Take 5 minutes and mentally review your list of reason to quit smoking. Remember how you felt when you decided to quit. [A3]
Be prepared for a rollercoaster ride of emotions: denial, anger, bargaining, depression, acceptance, and complacency. These are normal emotions for your body when going through withdrawal. [A4]

Figure 7: Teaching users: Emotion work.

People adjust quickly to feeling fitter. Remember that it's the fact you've quit that makes you feel so good. [A5]
Speak to yourself in a positive manner: 'I'm working on a healthy future', 'I'm doing well'... convince yourself to keep going. [A5]
You might sometimes glorify the past and miss smoking. This is nothing more than the 'rose-tinted spectacles' phenomenon! [A5]
The more you fight against difficult moments, the more intense they feel. The key to success lies in accepting and letting go. [A5]
Use your wildest imagination to exaggerate this moment, just like in a dramatic film. You'll soon see how ridiculous it is. [A5]
Humour helps to put moments of panic into perspective. Laughter reduces stress. Even anticipating a laughing fit ('this'll be fun!') can help! [A5]
Everything has its time, even your panic. Instead of being afraid of it: observe this feeling, it is something you can bear.[A5]

Figure 8: Teaching users: Emotion work in finer detail.

Exsmokers iCoach (App 5) differs from the other applications by engaging emotion work in more depth: smokers are oriented to observe processes of emotional adjustment to change, to avoid nostalgia, to use humor and 'wild imagination' to combat cravings, to observe panic and to actually talk to themselves - in a clearly observable instance of emotion work (see Figure 8).

4.1.6 Identity Work

Schwalbe & Mason-Schrock (1996) use the concept of 'identity work' to refer to the discursive actions through which people create identities for themselves and for others, by using, modifying, and creating categories and subcategories adequate to their interactional interests. We can observe how applications propose the identity category of 'ex-smoker', positioning it in conflict with the tobacco industry and in contrast with the category of 'smokers' (see Figure 9).

Don't let the government and the tobacco industry earn even one more cent from you. [A1]
On average, smokers die 7-8 years earlier than non-smokers. [A1]
Don't be afraid to look ridiculous. The world looks up to the brave ones, even if they don't succeed at first.[A5]
This is not buckling under pressure from others! You stop smoking for yourself. To feel good and to feel healthy. [A5]
You are no-smoker now, and you are strong enough to resist the craving to smoke. Believe in yourself. [A3]
Smokers are comparable to alcoholics and heroin addicts. You fool yourself, and tell yourself it's not that bad an addiction to have. Treat yourself like an alcoholic and never touch another cigarette again! [A4]
Don't you choose what you do? As long as you keep smoking, you're at the mercy of a cigarette. Do you smoke or are you being smoked? [A5]

Figure 9: Teaching users: Identity work.

The main identity feature of the ex-smoker identity is its *autonomy*, in contrast with smokers' addiction. This autonomy is also marked as freedom from others, from the government, and from tobacco companies.

An interesting aspect of identity work in smoking cessation applications consists in the proposed differentiation between users and their 'brains' (see Figure 10). Following a cognitivist vocabulary (Coulter 1979), applications instruct users to consider their brains as agents: thus, it is brains who learn, feel pleasure, or think – while

users are positioned as observers of these neuronal processes, gaining freedom from their biological determination.

> Throw away all cigarettes you have. You need to resist just a little longer. You are still getting used to non-smoking in certain situations. Your brain is still learning. [A2]
> You can train yourself to enjoy other things than smoking. Your brain's pleasure centres are currently overly focused on cigarettes as a source of pleasure. [A5]
> After quitting, your brain will be well supplied with blood, you will be better at thinking. [A1]

Figure 10: Teaching users: 'Observe your brain'.

4.1.7 Body Work

Applications commonly advise ex-smokers to engage in physical exercise and other forms of physical activity (sauna, brushing one's teeth, drinking water, raising sugar levels etc. – see Figure 11).

> Instead of smoking, go swimming. [A1]
> Get physical: Physical activity reduces craving. Put on shoes and go running. Do push-ups, knee-bends, and so on. [A2]
> When blood sugar levels drop, cravings can seem more powerful while you feel less able to manage them. Eat fruit (apple, grapes, kiwi) or a yogurt to feel better. [A3]
> Craving a cigarette? Brush your teeth and enjoy that fresh taste. [A3]
> Drink cranberry juice, it helps remove the poison nicotine from your system. [A4]
> Self-control issues could be down to a lack of sugar. To keep yourself in line: eat some fruit or bread. [A5]

Figure 11: Teaching users: Body work.

The autonomy-based identity work presented in section 4.1.6 explains applications' *reluctance to direct users towards medical or pharmacological help*. Among the 498 advice messages that we have analyzed, only one instructs users to "Seek help from a doctor or pharmacist" [A2]. The second mention of a doctor serves to enhance autonomy at the expense of medical authority: "Why are you considering quitting? The motivation to quit must come from you. Don't do it to please your family or your doctor" [A5].

Unlike other solutions, Exsmokers iCoach (App 5) includes a dedicated page with a detailed presentation of pharmacological alternatives available for ex-smokers, recommending users to seek medical help.

4.1.8 Environmental Work

Last but not least, there is substantial advice related to observing and intervening on one's material and social environment. The concept of 'smoking clues' refers to situational elements that elicit the need to smoke. Applications provide some advice to diminish these contextual clues, and to enhance material and relational support for the difficult transition (see Figure 12).

> Throw away everything that is associated with smoking, ashtrays, lighters, cigarettes, … [A1]
> Distract yourself: Call somebody. Write a text message. Go for a walk. Play a game on your phone. Make a coffee. [A2]
> Escape the situation: Go to some other place. Talk to some other people. Avoid trigger situations [A2]
> Keep pictures of loved ones on you. When you will feel cravings, look at those pictures and think about all the love you have for these people. [A3]
> Where are you right now? This location is probably strongly associated with smoking. Get away for a bit. Try to stop smoking in this spot. [A5]
> Don't get caught out by events that push you back to your old habits. You're doing the right thing: keep it that way! [A5]

Figure 12: Teaching users: Environmental work.

5 LEARNING FROM USERS

As we have discussed above, apps (and app developers) gather little information from users, mostly focused on their quantified history and patterns of smoking. This information serves to customize the quantified messages that position users on a short- and long-term trajectory from heavy health risks to non-smokers' health risks. Textual advice is not personalized – with the exception of App 5 that adjusts messages according to users' stage in the process of smoking cessation.

There are two main possible ways of personalizing communication with the user. On the one hand, apps could generate a more detailed user profile, estimating the motives of smoking, the degree of addiction, and the severity of withdrawal symptoms, which is associated with the probability of relapse. These profiles can be constructed using available online scales, as we discuss in the next section. On the other hand, apps could ask users to rate the received advice and tips, and infer a profile from users' expressed preferences – that may vary as regards the domain of action (perceptions, emotions, body, or environment), tone (more or less focus on

fear from smoking damage, more or less acceptance of *guilt*, more or less *humour*) as well as a preference for shorter or longer messages, or reflective versus action-oriented tips.

6 LEARNING FROM ONLINE SCIENTIFIC LITERATURE

Our analysis supports Abroms' et al. (2011) conclusion that smoking cessation apps do not adhere to medical advice: they largely rely on a common-sense 'cold-turkey' model and ignore the helpful potential of counselling and pharmacotherapy.

There is no shortage of scientific literature on smoking and smoking cessation freely available online. We propose a brief inventory as a guide for developers in reviewing the literature.

Medical research is useful for understanding the benefits and often the necessity for nicotine replacement therapy (Benowitz 2008).

Psychological scales are relevant for better comprehending the variety in smoking motives, addiction, and withdrawal symptoms and intensity:

- Scales for assessing the *intensity of addiction to nicotine*: the Fagerstrom Test for Nicotine Dependence (Fagerstorm et al., 1990); The Cigarette Dependence Scale (Etter et al., 2003); the Nicotine Dependence Syndrome Scale (Shiffman et al., 2004); the Autonomy over Tobacco Scale (Wellman et al., 2011);
- Scales for identifying users' *motives for smoking*: The Wisconsin Inventory of Smoking Dependence Motives: WISDM 68 (Piper et al., 2004) and WISDM 37 (Smith et al., 2010);
- Scales for assessing users' withdrawal: the Wisconsin Smoking Withdrawal Scale (Welsch et al., 1999) and the Wisconsin Predicting Patients' Relapse Scale (Bolt et al., 2009).

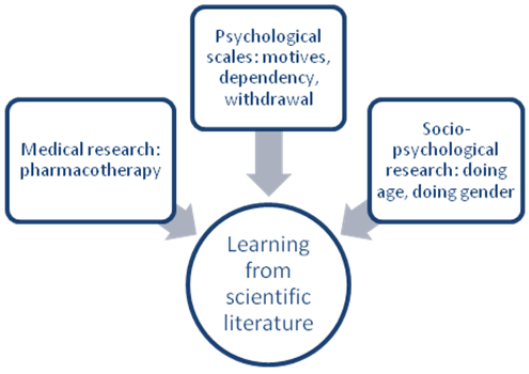


Figure 13: Domains of scientific literature on smoking.

Apps advice does not actually address the functions of smoking in users' daily life, as detailed in the WIDSM scale, for example. Most tips only refer to cravings and loss of control; some refer to environmental cues. Taking into account the other motives for smoking may improve the subjective relevance of advice.

While social cues feature as a distinctive motive for smoking in psychological scales, the implicit social process remains rather abstract. The concrete social uses of smoking can be better understood through socio-psychological research that points to the *identity work that smokers achieve through smoking*. Smoking is often important for expressing one's gender, class, and age identities (Rugkasa et al. 2003; Barbeau et al. 2004; Amos & Bostock 2007) – a feature easily visible in smoking adds that have for long framed the cigarette as a symbol of masculinity and, respectively, femininity, in conjunction with age and social class. Apps could address the usefulness of smoking for gender and age identity work through their advice, acknowledging this function and proposing replacements.

7 CONCLUSIONS

Apps *teach users* new ways of perceiving the world and their bodies, of dealing with their emotions, of acting on their body and environment.

Still, apps learn surprisingly little *from users*: user profiles are summarily sketched to serve the quantification modules of health predictions, and there is no customization of advice and tips.

There is also little reliance on *scientific literature*. Mobile apps propose a user identity based on the core value of autonomy and the lay theory of 'cold-turkey' quitting. Users are represented as 'lone riders', in a difficult quest, assisted only by their coach – the application, and other peers (family, friends, networks). This identity project makes applications reject medical and pharmacological advice, in contrast with evidence-based recommendations for smoking cessation. Beyond medical literature, we propose that psychological research concerning scales of smoking addiction, motives, and withdrawal, and socio-psychological research on identity work through smoking offer valuable and easily accessible resources for app developers in order to create rich user profiles, develop their repertoires of advice, and personalize tips to users' preferences and needs.

ACKNOWLEDGEMENTS

This article has been supported by the project "Sociological imagination and disciplinary orientation in applied social research", with the ANCS/UEFISCDI grant PN-II-RU-TE-2011-3-0143.

REFERENCES

Abroms, L. et al. 2011. iPhone Apps for Smoking Cessation: A Content Analysis. *American Journal of Preventive Medicine*, 40(3), 279–285.

Amos, A. & Bostock, Y. 2007. Young people, smoking and gender A qualitative exploration. *Health education research*, 22(6), 770–81. http://her.oxfordjournals.org/content/22/6/770.long.

Barbeau, E. M., Leavy Sperounis, A. & Balbach, E. D. 2004. Smoking, social class, and gender: what can public health learn from the tobacco industry about disparities in smoking? *Tobacco Control*, 13(2), 115–120. http://tobaccocontrol.bmj.com/content/13/2/115.full.

Benowitz, N. L. 2008. Neurobiology of Nicotine Addiction: Implications for Smoking Cessation Treatment. *The American Journal of Medicine*, 121, 3–10.

BinDihm, N., Freeman, B. & Trevana, L. 2012. Pro smoking apps for smartphones: the latest vehicle for the tobacco industry? *Tobacco Control*, 1–7.

Bolt, D.M. et al. 2009. The Wisconsin Predicting Patients' Relapse questionnaire. *Nicotine & Tobacco Research*, 11(5), 481–492. http://www.ctri.wisc.edu/Researchers/WI PREPARE09.pdf.

Cohn, A. et al. 2011. Promoting Behavior Change from Alcohol Use through Mobile Technology: The Future of Ecological Momentary Assessment. *Alcoholism: Clinical and Experimental Research*, 35(12), 2209–2215.

Coulter, J. 1979. The Brain as Agent. *Human Studies*, 2, 335–348.

Etter, J. F., Le Houezec, J. & Perneger, T. 2003. A Self Administered Questionnaire to Measure Dependence on Cigarettes: The Cigarette Dependence Scale. *Neuropsychopharmacology*, 28, 359–370. http://knowledgetranslation.ca/sysrev/articles/project21/Ref ID3219 20090722180832.pdf.

Fagerstorm, K. O., Heatherton, T. F. & Kozlowski, L. T. 1990. Nicotine Addiction and Its Assessment. *Ear, Nose and Throat Journal*, 69(11), 763–765. http://www.dartmouth.edu/~thlab/pubs/90_Fagerstrom_etal_ENTJ.pdf.

Flaherty, M. G. 2003. Time Work: Customizing Temporal Experience. *Social Psychology Quarterly*, 66(1), 17–33.

Frederick, S., Loewenstein, G. & O'Donoghue, T. 2003. Time Discounting and Time Preference: A Critical Review. In G. Loewenstein, D. Read, & R. Baumeister, eds. *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*. Russell Sage Foundation, 13–86.

Hochschild, A. R. 1979. Emotion Work, Feeling Rules, and Social Structure. *American Journal of Sociology*, 85(3), 551–575.

Piper, M. E. et al., 2004. A multiple motives approach to tobacco dependence: The Wisconsin Inventory of Smoking Dependence Motives (WISDM 68*). Journal of Consulting and Clinical Psychology*, 72(2), 139–154. http://www.ctri.wisc.edu/Researchers/AcceptedWISDMManuscript.pdf.

Rugkasa, J. et al. 2003. Hard boys, attractive girls: expressions of gender in young people's conversations on smoking in Northern Ireland. *Health Promotion International*, 18(4), 307–314.

Schwalbe, M. L. & Mason Schrock, D., 1996. Identity work as group process. *Advances in Group Processes*, 13, 113–147.

Shiffman, S., Waters, A.J. & Hickcox, M. 2004. The Nicotine Dependence Syndrome Scale: A multidimensional measure of nicotine dependence. *Nicotine & Tobacco Research*, 6(2), 327–348. http://www.researchgate.net/profile/Saul_Shiffman2/.

Smith, S. S. et al. 2010. Development of the Brief Wisconsin Inventory of Smoking Dependence Motives. *Nicotine & Tobacco Research*, 12(5), 489–499. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2861888/pdf/ntq032.pdf.

Wellman, R. J. et al. 2011. Psychometric Properties of the Autonomy over Tobacco Scale in German. European *Addiction Research*, 18, 76–82. https://umassmed.edu/uploadedFiles/fmch/Research/Publications/Psychometric properties article.pdf.

Welsch, S. K. et al. 1999. Development and Validation of the Wisconsin Smoking Withdrawal Scale. *Experimental and Clinical Psychopharmacology*, 7(4), 354–361. http://www.ctri.wisc.edu/Researchers/WSWS.pdf.

3D Online Virtual Museum as e-Learning Tool
A Mixed Reality Experience

Dragoş Gheorghiu[1] and Livia Ştefan[2]
[1]*Doctoral School, National University of Arts, 19 Budişteanu Str., Bucharest, Romania*
[2]*Department of Computing Applications, Institute for Computers, Calea Floreasca 167, Bucharest, Romania*
gheorghiu_dragos@yahoo.com, livia.stefan@itc.ro

Keywords: 3D Virtual Museum, e-Learning, Virtual Reality, Mixed-Reality.

Abstract: The paper discusses the use of 3D online virtual museums as educational instruments for children in economically disadvantaged communities. It discusses the advantages of simulation and immersion as learning enhancements, compared to traditional learning methods, using a case study based on the currently on-going work of development of a virtual museum (www.timemaps.net/timemap/?page_id=6). Presenting chronologically layered information, the museum offers a series of e-learning modules teaching traditional crafts technologies. The current audience is represented by children from a Romanian village community. The paper covers a description of the methodology used, as well as an evaluation of the usefulness of the solution by the target audience. In conclusion, the paper outlines a strategy for further development of the proposed e-learning solution.

1 INTRODUCTION

Distance-learning (Aldrich, 2009:48) has become a method of choice, using virtual environments in order to provide accessible education in underprivileged environments. We have used it as a solution to facilitate the access to education in remote and economically disadvantaged communities, by initiating an educational project – "The Maps of Time - Real communities-Virtual Worlds-Experimented Pasts" (Grant PN II IDEI), linking these communities to the research centres and urban education institutes, such as the National University of Arts Bucharest (NUA).

The paper presents one of the educational experiments within this research project, conducted by Professor Dragoş Gheorghiu in Vădastra village, Oltenia County, in southern Romania. Participation in the experiment has included specialists and teachers from our university, as well as teachers and children from three secondary schools. The experiment consisted in the implementation of a 3D virtual museum comprising 3 archaeological levels, with realistic reconstructions of ancient objects and using Mixed Reality (Milgram and Kishno, 1994) to augment the 3D virtual environment with video films. The educational content (3D reconstructions and the professional video films), was developed by masters' program students and artists from our university, under the supervision of Professor Dragoş Gheorghiu.

The objective was to create e-learning solutions supporting the teaching of traditional technologies in an intuitive, engaging, play-immersive manner, leading to the enhancement of the traditional, curriculum-based learning, through the use of innovative and cost-effective ICT technologies, and learning by doing (Dewey, 1938; Schank, 1995). Finally, the experiment aimed at the creation of a learning community. The paper presents the details of the learning experiment as a case study, in relation with other experiments that were performed in similar international and national projects, followed by an evaluation of the learning outcome. The conclusions of the experiments are presented,



Figure 1: The entrance in the virtual museum.

validating the use of the 3D virtual museums as modern and effective learning instruments. The paper closes with an outline of the future development strategy.

2 THE IMPORTANCE OF MUSEUMS FOR EDUCATION

The limitations of traditional learning methods have been addressed in several papers. Relan and Gillani (1997) characterize formation using the traditional educational settlements as follows: a) Spatial and temporary structures, in which learners should adapt; b) Learning geographically limited to classrooms, laboratories, etc. and disciplines are taught at predetermined times and in a pre-established order; c) The physical presence of student and teacher in the same classroom is an indispensable requirement, so that learning can occur (see also Campas, 2008).

2.1 Traditional Museums

As early as the 19th century the museums started to be reconsidered for a systematic use as educational instruments. These "were placed on the front line in the educational agendas of the period. This was evident in the increasingly close connections between museums and the development of compulsory public schooling, and in the stress that was placed on the value of museums as instruments of adult education" (Bennet, 2004).

On the other hand, the educational results to be achieved from visiting traditional real-life museums have their limitations, due to: a) relatively conventional presentation methods; b) limited accessibility (in time and space); c) limited visiting time; d) static, non-interactive visits, often without a guide to provide explanations.

2.2 Modern Museums

As Germain Bazin anticipated long time ago, we live now in "The Museum Age" (de Leeuw 1999: 79). In the 21st century the modern museums improved the users' experience through the introduction of interactivity, and of the online versions of the institutions. Beside modern ICT technologies, today's museums promote:

- a new way to look at the past (Wheat-Stranahan et al., 2007);
- online multimedia content; The European Virtual Museum, implements a dynamic web site based on "user-friendly interactive databases", stating that a "System of Management for interactive databases (RDBMS) is crucial in order to create a virtual museum";
- an immersive experience using Virtual Reality applications and 3D modelling in online museum environments (Hermon, 2008).

2.3 The Advantages of Virtual Museums

Several advantages can be derived from creating and visiting virtual museums:

- unrestricted accessibility regardless of time and location;
- the capacity to display ancient monuments and objects in their complete form;
- (virtual) access to otherwise (physically) inaccessible places in a monument;
- the capacity to interact with the museum's objects;
- the capacity to enable virtual tourism (Maiorescu and Sabou, 2013; El Hakim et al., 2004);
- the capacity to use simulation and recreation as educational tools (Politis and Marras, 2008);
- the capacity to access complex museum presentations as reconstructed objects alone, or in their historical contexts, as well as objects augmented with explanations;
- an educational exploration at reduced cost, requiring only a good Internet connection;
- interactive communication of cultural heritage in public spaces (Trapp et al., 2012).

2.4 The Advantages of Learning in Online 3D Virtual Environments

The advantages of learning in 3D virtual environments, as a form of distance education (Annetta et al., 2010), have been outlined by several theoretical papers. Winn (2002) states that one important ability of 3D virtual environments is reification, by which objects that are not accessible to the human senses are transformed into "visible, audible and tangible objects" (Gell-Mann, 1994).

Another characteristic of 3D virtual environments is user's immersion in the virtual space, which creates a feeling of presence and stimulates the learning outcomes (Winn, 1995; see also Dickey, 2003; Dede, 2009). The immersion in an artificial environment can help learners to better understand the tri-dimensional dynamic processes (Winn, 1995), facilitates both the verbal and non-verbal communication and collaboration (Kappe and Gütl, 2009) or other "important life skills" (Annetta

et al., 2010).

The physical presence has a strong cognitive function, leading to a higher level of involvement of the learners in the learning process. This implication is considered more important for the "conceptual modification than the cerebral activity, typically necessary for problem solving based learning" (Clark, 1999).

A virtual world can enhance an experiential learning process through activities such as simulation and role-play. It also "promotes immersive learning as the learner can potentially experience the emotions and thoughts of someone in a simulated situation" (Macedo and Morgado, 2009).

These environments also support the self-paced learning style, i.e., the possibility for the student to resume or repeat the learning process.

For human sciences, 3D environments can have a stronger cognitive impact, as they give learners the feeling of being in the simulated reality, thus improving the understanding and memorization of the information. Possessing an important spatial component, the 3D virtual environments allow users to freely change the visualization perspective. Finally, not the least important aspect, 3D virtual environments allow the simulation of virtual time travel.

3 THE MIXED REALITY PARADIGM

According to the definition of Milgram and Kishino (1994) Mixed Reality (MR) is a "subclass of VR related technologies that involve merging of real and virtual worlds".

Augmented Reality (AR) is a technology, which allows an enhancement of the reality using virtual information (Azuma et al., 2001). When a virtual environment is augmented with content from the real world, the process is called Augmented Virtuality (AV). Milgram and Kishino (1994) proposed an integration of the two technologies, called Mixed Reality.

The augmentations enhance the learning outcome since "the AR/MR systems allow the augmentation of the user's perception of the real world, but also the use of existent visual and spatial abilities and an enhancement of the interaction capacities of the users" (Holza et al., 2011). This is why the AR systems have to "give their users the illusion that digital objects are in the same space as physical ones" (Costanza et al., 2009), and therefore "the digital objects need to be precisely positioned into the real environment and aligned with the real objects in real time" (Azuma et al., 2001). When the augmentation is applied to virtual environments this requirement is not imperative.

4 SIMILAR WORK

There are currently several projects which address the presentation of the virtual heritage: "Museums are now taking a much wider interest in the use of ICT to leverage better value out of their collections; a simple internet search on museums and ICT reveals many hits detailing museum activity "(White et al., 2004).

It is worth mentioning the online versions of the Romanian Peasant Museum (Bucharest), which offers virtual tours and online visualizations' or the Louvre Museum, which exposes 3D reconstructions of the museum's artifacts. Even though they are not implemented with an explicit educational objective, they are significant examples of the innovative technologies employed by online virtual museums.

Recently, several EU-funded research projects have aimed at developing innovative technologies for museums to create interactive 3D Virtual Exhibitions on the Web. For instance, AMIRE is a EU IST Program dedicated to the efficient creation and modification of MR applications, authoring metaphors and generic design recommendations and procedures. In the ARCO project the elaboration process is described as follows: "virtual exhibitions are created by digitising museum artefacts, which are then transformed into Virtual Representations, which can be X3D or VRML models or scenes" (White et al., 2004).

Another example is the virtual project "A History of the World in 100 objects" (AHOW) run by BBC Radio 4 and the British Museum in 2010 (Cmeciu and Cmeciu, 2013).

Within our project we sought to experiment with several e-learning solutions for the community of teachers and children at the secondary school of Vădastra, in order to address the needs for informal learning using modern IT technologies and e-learning paradigms, and evaluate the effectiveness of each particular solution for long-term use; for this purpose we developed mobile-learning solutions described in Ştefan and Gheorghiu (2013) and Gheorghiu et al. (2013).

5 THE RATIONALE OF THE RESEARCH PROJECT

The village of Vădastra, Olt county, southern Romania, was selected for several reasons: it is a clear example of the deep cultural crisis experienced by rural peripheral settlements in the early 21st century, a crisis brought about by rapid social changes; it is here that a decade ago our university had initiated a project to preserve the immaterial heritage of the village traditional technologies.

We felt that an overview of the history of the place (a palimpsest of successive habitations from prehistory to modern times) could, apart from facilitating learning, recreate a sense of its identity, which had been erased during the country's totalitarian period. As a topic that might interest the community through its pragmatic, economical aspect, we insisted on the rediscovery of ancient and traditional technologies, which form the immaterial heritage of the place.

Centering the learning strategy on this subject, the artists and technicians from our university began to develop e-learning lessons teaching different crafts representative for this geographical area.

Each of the solutions tested in the research project has proven its effectiveness in enhancing learning especially when solutions were designed with a clear learning outcome and also understandable for children to use, but also revealed some limitations: the difficulty to use the solutions outside the project because it required prohibitively expensive equipment; insufficient teacher support and insufficient engagement in the learning process.

This led us to the development of a more intuitive solution, usable with the technological means that the school already had, or could afford. Initially, we designed a 3D virtual museum as an informal e-learning environment for ancient traditional technologies, considering that this would better serve our learning objectives and prove appropriate for the target community group. Nevertheless we mixed the learning in a virtual environment with practical classes in a blended learning perspective (i.e., mixing several technologies and learning styles for an optimal learning outcome).

5.1 Research Questions

Our archaeological and pedagogical strategy was formulated after a research which can be condensed in the following questions:

a) which are the historical stages that can define the identity of a place?
b) which are the contexts representative for each of the identified periods?
c) which are the technologies representative for each of the chosen contexts?
d) which are the objects representative for each of the technologies?

6 CASE STUDY: THE 3D VIRTUAL MUSEUM FOR A VILLAGE COMMUNITY (VĂDASTRA, SOUTH OF ROMANIA)

After several archaeological experiments, a database of videos and 3D reconstruction was created and explored as a 3D virtual museum. Designed by Professor Dragoş Gheorghiu on 3 levels, to suggest the stratigraphy (sequence) of the major historical periods in the Vădastra village (i.e. the modern, Roman, and prehistoric), the museum contains exhibits specific to each period, which are augmented with videos representing craft technologies.

These augmentations of the virtual space are embedded in 3D virtual panels and are displayed as still picture frames. A click on the panel launches the videos in a separate window player.

The 3D virtual museum is implemented as a Unity3D application integrated in the project's website (www.timemaps.net/timemap/?page_id=6).

The virtual museum can be explored in a game-like manner: the mouse is used for camera (perspective) control; the direction keys, for spatial - temporal navigation.

To resemble a real museum, a character called "The Guide" was created, with the role of providing an explanation at the entrance of each level of the museum. "The Guide" is designed as a friendly persona, to engage the users, both children and teachers, to explore the museum. It is similar to a *Non-Player Character* in certain video games, but designed with a limited functionality within the virtual museum application.

6.1 The First Level of the Museum

The virtual Museum opens with a first "artistic" room: here, visual artists and experimentalists' interpretations of ancient and traditional objects are presented, as models of revitalization of both the high culture art and the folk art. Children learn that

modern art can be inspired by the traditional one, and artistic techniques to manufacture various objects are explained to them.

From here on children travel in space and time into the second room of the museum, which exhibits Roman objects, reconstructions of the fragments found in the village and in the nearby region and exhibited at the Romanaţi Museum in Caracal.

The time travel continues to the prehistoric past into the room displaying a Chalcolithic house (excavated in Vădastra in the 70s), together with reconstructed prehistoric everyday life objects.



Figure 2: The "Guide".

6.2 The Second Level of the Museum

The second level is represented by the Roman period. Many technologies, such as manufacturing and pottery, or the textile art, have survived the period with very small changes until the pre-industrial period. This level has been reconstructed to emphasize a *villa rustica*, discovered by the first author near the village, in which different domestic industries are showcased: pottery, weaving, glass manufacturing or metal casting.

All the tools necessary for those technologies, the complete form of the villa and the surrounding landscape were reconstructed under the form of a Unity 3D Virtual Reality. The hyper-realistic reconstruction allows, as in the previous case, a strong immersion of the visitor. The augmentation of the 3D virtual space with records of the technologies to be studied, contribute to this process of immersion.

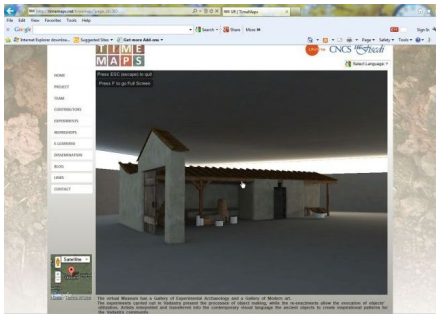


Figure 3: The Roman room of the virtual museum.

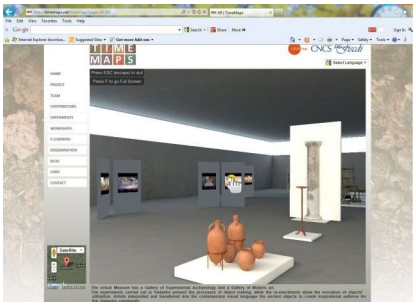


Figure 4: Objects in the Roman room of the virtual museum.

6.3 The Third Level of the Museum

The third educational level of the museum is represented by the Chalcolithic village (5th millennium BC), belonging to the eponymous Vădastra culture. Here also we wanted to highlight the technological traditions, especially the weaving one, which, as has been noted after the exceptional archaeological discovery of a prehistoric textile at Celei, near the Vădastra village, had continuity until relatively recent times. The children experience the living space of the prehistoric settlement and access the augmentations of the reconstructed house under

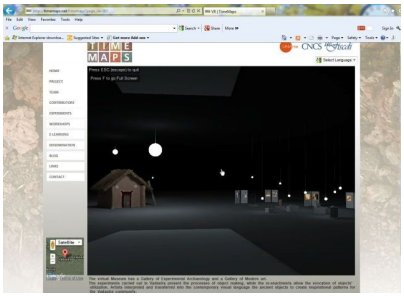


Figure 5: The prehistoric room of the virtual museum.

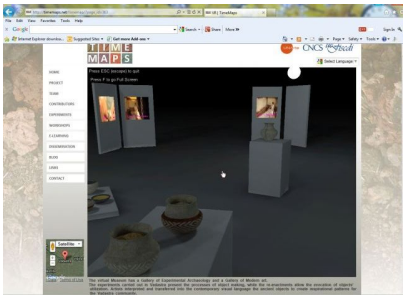


Figure 6: Objects in the prehistoric room of the virtual museum.

the form of videos presenting house construction techniques and weaving techniques for the Celei textile.

6.4 Technical Description

The modelling of the virtual museum was performed with Autodesk 3ds Max and Maya specialized software, which allowed not only to achieve the realistic reconstructions, but also to implement optimization techniques for the real time rendering of the 3D scenes and objects, such as texture baking and LOD (levels of details) optimization. The Unity3D platform was used in order to integrate the virtual museum in an online environment and to implement the Mixed-Reality techniques.

Unity3D is an authoring tool and a gaming engine for 3D video games and 3D interactive content, which support programming in C ++ or C # and also provides a powerful scripting language.

The reasons for selecting this platform were: performance optimization for real time rendering; multi-platform support for the Unity3D web plugin (Microsoft Windows and Mac OS X) and for applications (Windows, Mac, Xbox 360, PlayStation 3, Wii, iPad, iPhone, Android); and the advanced graphical editor.

To run the application in the web browser, the Unity3D web plugin is required. This can be a disadvantage when repeatedly running the application, as the web plugin has to be loaded every time. We tried to reduce this inconvenience by capturing the user's attention with explanatory texts.

6.5 The Methodology

The main purpose of our Virtual Museum is to integrate pedagogy and new technology (Annetta et al. 2006). The Museum offers a presentation of local history and of the technologies specific to each historical period in their historical context, by using virtual reconstructions. By presenting historical contexts virtually reconstructed, the design of the Virtual Museum offers the learners a structured information on: a) the objects of each context, in a 3D reconstruction; b) videos with technological gestures for creation of these objects; c) a series of performances of the usage of these objects during the historical periods.

For supporting the e-learning tool, each technology was filmed from at least two view points: the actor's view and the observer's view. In this way, the child who wishes to learn a certain technology has a complete understanding of the technological gestures. Teaching in a virtual and immersive environment allows flexibility in structuring the course material; the teacher having the following pedagogical objectives: a) acquisition of historical data; b) acquisition of technical data and skill development during the presentation of the technologies (Annetta et al. 2006).

It is worth mentioning the exploratory aspect of the pedagogical and learning activities, by engaging and capturing the attention of the children involved in the learning process. For this aspect we used an approach focused more on the direct teacher-learner relation. First, during the history or technological lessons, the local teacher navigates, together with the children, through the site and the virtual reconstructions, explaining to them the reconstructed local history. The second phase is the professional/practical approach: we organized during the summer of 2012 and 2013 a series of workshops (on textile and glass engraving) for the children in Vădastra, where they worked under the coordination of a teacher or a technician. During the school semester, this experience was evaluated remotely, from our university, in this way: a) the teachers transmitted to the children a theme from the Virtual Museum, which they individually studied. b) the teachers organized video lessons in which each child presented the results of his/her work, which has been further corrected and evaluated. In this case of learning by doing, the evaluation through questionnaires was not applicable.

Beside the 3D Virtual Museum, traditional methods of distance education were also used. Skype-based videoconferences were organized at our university in order to connect the academic specialists with the teachers and children from Vădastra, and to offer them various lesson plans. For example, we proposed to visit the Neolithic village and investigate how prehistoric art can be transformed into modern art.

At the beginning the university artists proposed different learning paths and explained in detail each object and its corresponding technology. Thus, using guided tours via Skype we could coordinate the actions of the school teachers to efficiently use the 3D Virtual Museum, eventually leading to sufficient proficiency so as to be able to create their own lessons. As a conclusion of these learning practices, Vădastra School has introduced two applied classes (on ceramics and textile), where the children could reproduce many of the Virtual Museum's objects. Consequently, our university has initiated with our students an online course in order to supplement the activity of the village teachers.

7 RESULTS EVALUATION

In the case of using the Virtual Museum for learning about the local history and for revitalization of arts and crafts of immaterial heritage, we applied two methods: a) the one of the local history teacher who will use semi-structured interviews; b) the one of the technology teacher who will use the visual contact with the children (direct or via Skype) for evaluation.

The 3D virtual museum was experimented for a six-month period, during which the project team provided technical and pedagogical assistance. The experiments took place in the Vădastra secondary school community, grades 1-4, 5-8, as well as in two urban school communities, Râmnicu Vâlcea (Central Romania) and Piatra Neamţ (North Romania), allowing us to verify the effectiveness of our learning methods in different cultural environments.

Throughout the experiments, monitoring was performed by collecting feedback, suggestions, and recording of technical problems, thus gathering an important quantity of useful information. The data was collected using Google Docs together with the answers from 6 teachers (aged 25-40) and 9 children (aged 8-15).

To use the site of the project as an applied learning instrument, we further proposed a set of refined questions focused on the virtual and real experience of the children, which starts from mental immersion and ends with the manufactured objects:

Table 1: Questionnaire for results evaluation.

Which was the most interesting historical period?
Which were the most interesting objects?
Which object have you chosen to make?
How much time did you work for the practical workshop?
How much time did you work with the online teachers?
How realistic did you find the architectural 3D reconstructions? (answer with a mark from 1 to 7)
How realistic did you find the objects' 3D reconstructions? (answer with a mark from 1 to 7)
What have you learned from the 3D experience (local history, manufacturing of objects, use of objects, a.o.)?
Which were the criteria for choosing an object (form, texture, size, position in context)?
How many times did you play the film with the 3D reconstruction of your object?
What have you learned from these films (form construction, decoration making, a.o.)?
What can you make in the present? Can you manufacture an object without help? (answer by enumeration of the technological operations)

We also proposed a peer assessment (Aldrich, 2009), through which we asked the Vădastra school teachers to assess the professors who coordinated the lessons.

8 DISCUSSION

The post completion analysis of the project identified the following limitations of the approach: in some cases establishing the logical connection between the 3D reconstructed objects and contexts and the technological film required the viewers (children) to have prior knowledge of the subject, or to receive explicit guidance. Therefore, the learning process within a virtual environment must allow the teachers to design by themselves the guided tours, i.e. to have a greater degree of flexibility.

In spite of the opinion expressed in Di Blas et al. (2008), we considered the educational importance of 3D realistic reconstructions for the museum objects while we performed simplifications regarding the modelling of the museum space.

The technical problems were related only to the quality of the Internet connection and the delay caused by the loading of the Unity3D web plugin at each application launch.

We used some of Aldrich's (2009) suggestions to perform assessments for "grading student performance", namely to write a paper about the experience in the virtual museum, to keep a journal during the experience, and create a multimedia production, in our case weblog and Facebook postings.

We appreciate that the educational advantages of the mixed-reality experience in an immersive 3D environment was highlighted by the following:

- The visual presentation of the exhibits corresponding to different historical periods allowed a good message retention and understanding.
- The videos have increased the educational value of the learning experience since they improved the viewer's sense of reality.
- The return to the 3D environment after viewing the movies required a certain discipline and patience in exploring the museum.
- After the guided tours, the children were able to perform self-directed learning.
- We have identified cognitive (concentration, memorization) and affective (enthusiasm, a sense competition in navigation the space) components to the learning process.
- The computer literacy was enhanced with

concepts necessary to understand the new learning environment (e.g. Virtual Reality).

9 CONCLUSIONS

The mixed-reality experiment presented was designed and modelled with a well-defined educational objective. The 3D virtual museum has proved to be a creative tool, used both by teachers and children in a flexible manner that allowed learning without the constraints of the traditional learning process and in an immersive and playful manner. The 3D mixed-reality experience has allowed children to immerse in a spatial and temporal virtual environment while still maintaining the connection to real life by watching the video recordings.

This experiment led us to the conclusion that 3D virtual museums, as an intuitive exploration space, are a valid instrument of education. A 3D online virtual museum designed for education can support learning by doing and can accommodate a large and distributed learning community. The experiment showed interest and enthusiasm from the children and teachers for the novel learning environment, as a validation of the technology and approach we used.

10 FURTHER DEVELOPMENTS

In order to capitalize on the present experience we decided to develop a collaborative version of the 3D virtual museum: Vădastra participants from Bucharest, Râmnicu Vâlcea (Central Romania) or Piatra Neamț (North Romania) will have the possibility of entering the museum in real time.

The new museum will allow the teaching of traditional technologies under the form of guided tours, accompanied by explanations. The function of the "Guide" avatar will be extended providing a more complex and animated character that will deliver explanatory texts in front of each object and indicating for each object the technological video which needs to be played; the avatar movements will be determined by the teacher: for example, a professor in a remote location (our university or Vădastra school) will indicate course keywords, such as "Roman pottery" and the avatar will guide the users only on the Roman level, presenting the ceramic objects and the technology of pottery making. The children will follow each object presentation and will be able to ask questions using a chat component. At the end of each lesson, children will answer the questionnaire for learning assessment using the qualitative and semi-structured interviews presented at the Methodology paragraph (see Table 1). The children will be also evaluated by their objects and sketches made during the lessons. The learning method that we propose is usable for skills development in manufacturing of objects with low level of difficulty, but at the same time it offers the theoretical learning of making more complex objects. The future strategy mentioned above will require the design of the roles for each category of users (teachers, children), and the activation of the networking component of the Unity3D platform in order to support the development of a Multiple User Virtual Environment (MUVE) interactive application. This will also require a more powerful hardware infrastructure to ensure the required real-time performance.

The project presented will help to create a "community of learners" (Veen and Vrakking 2006). An interesting case to be discussed is how our approach generalizes in different European communities. During 2013, we presented the project to two urban communities from Portugal (Mação and Abrantes), which proved interested to use the results of the project as a pedagogical instrument for protection of local immaterial heritage. Recently, other communities from UK and the Netherlands also adhered to the project.

The 3D Virtual Museum will therefore support an extended virtual learning community, with numerous real time participants and mixed-reality lessons.

ACKNOWLEDGEMENTS

The authors wish to thank the principal of the Vădastra school, Vădastra village, Olt County, Romania, and all the teachers and children for their help and involvement in our experiments. The authors also thank the professors and technicians from our university for their collaboration.

Last, but not least, many thanks to Mr. Bogdan Căpruciu for reviewing the English text and to the anonymous reviewers whose comments helped us to improve the text.

REFERENCES

Aldrich, C., 2009. *Learning online with games, simulations and virtual worlds. Strategies for online instruction*, Jossey-Bass. San Francisco.

Annetta, L. A., Murray, M. R., Laird, S. G., Bohr, S. C., and Park, J. C., 2006. Serious Games: Incorporating Video Games in the Classroom, In *Educause Quarterly* 3, pp. 16-22.

Annetta, L. A., Folta, E., and Klesath, M., 2010. Use of Virtual Learning Environments in Distance Education, pp. 35-56. In *V-Learning - Distance Education in the 21st Century Through 3D Virtual Learning Environments*, Springer Netherlands.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B., 2001. Recent advances in augmented reality. In *IEEE Comput. Graph. Appl*. 21(6), pp. 34–47.

Bennet, T., 2004. *Past beyond memory, Evolution, Museums, Colonization*, Routledge. London and New York.

Campas, J., 2008. Teaching and Learning in Virtual Environments, pp. 48-63. In Politis, D., (ed.), In *E-Learning Methodologies and Computer Applications in Archaeology*, Information science reference. New York.

Clark, A., 1999. *Being there: Putting brain, body and world together again*. MIT Press. Cambridge, MA.

Cmeciu, C., and Cmeciu, D., 2013. Virtual museums – non-formal teaching e-civilization culture. In *Proceedings of The 9th International Scientific Conference eLearning and Software for Education,* Bucharest, April 25-26, 2013.

Costanza, E., Kunz, A., and Fjeld M., 2009. Mixed Reality: A Survey. pp. 47-68. In *Human machine interaction*, LNCS 5440, Springer.

Dede, C., 2009. Immersive Interfaces for Engagement and Learning. *Science Magazine* 323(5910), pp. 66-69.

de Leeuw, R., 1999. The History of Art between the 20th and the 21st Century. In *Diogenes*, No. 185, Vol. 47/1, Blackwell Publishers Ltd.

Dewey, J., 1938. *Experience and Education*, Kappa Delta Pi. New York.

Di Blas, N., Paolini, P., and Poggi, C., 2008. A Virtual Museum Where Students can Learn, pp. 388-406. In Politis, D. (ed.), *E-Learning Methodologies and Computer Applications in Archaeology*, Information science reference. New York.

Dickey, D. M., 2003. Teaching in 3D: Pedagogical Affordances and Constraints of 3D Virtual Worlds for Synchronous Distance Learning. In *Distance Education*, 24(1).

El-Hakim, S. F., Beraldin, J. A., Picard, M., and Godin, G., 2004. Detailed 3D reconstruction of large scale heritage sites with integrated techniques. In *IEEE Computer Graphics and Application*. 24(3), pp. 21-29.

Gell-Mann, M., 1994. *The quark and the jaguar: Adventures in the simple and the complex*. Freeman. NewYork.

Hermon, S., 2008. 3D Modelling and Virtual Reality for the Archaeological Research and Museum Communication of Cultural Heritage, pp. 57-72. In Oberländer-Târnoveanu, I. (ed.), *Museum and the Internet Presenting Cultural Heritage Resources On-line*, Archaeolingua. Budapest.

Holza, T., Campbella, A. G., O'Harea, G. M. P., Stafford, J. W., Martin A., and Dragonea M., 2011. MiRA - Mixed Reality Agents. In *International Journal of Human-Computer Studies*, Elsevier, 69 (4), pp. 251-268.

Kappe, F., and Gütl, C., 2009. Enhancements of the realXtend framework to build a Virtual Conference Room for Knowledge Transfer and Learning Purposes. In *EDMEDIA 2009.*

Macedo, A., and Morgado, L., 2009. Learning to teach in Second Life. In *EDEN Open Classroom Conference*, Porto, Portugal.

Maiorescu, I., and Sabou, G. C., 2013. Learning about heritage through Augmented Reality games, In *Proceedings of The 9th International Scientific Conference eLearning and Software for Education*, Bucharest, April 25-26, 2013.

Milgram, P., and Kishino, F., 1994. A Taxonomy of Mixed Reality Visual Displays. In *IEICE Transactions on Information Systems*, pp. 1321–1329.

Politis, D., and Marras, I., 2008. The Use of Virtual Museums, Simulations, and Recreations as Educational Tools, pp. 157-198. In Politis, D. (ed.), *E-Learning Methodologies and Computer Applications in Archaeology*, Information science reference. New York.

Relan, T. I., and Gillani, B. B., 1997. Web page based instruction and the traditional classroom: Similarities and differences. In Khan, B. H. (ed.), In *Web page-based instruction* Educational Technology Publications. New Jersey.

Schank, R. C., 1995. What We Learn When We Learn by Doing, Institute for the Learning Sciences Northwestern University, In *Technical Report* No. 60.

Ştefan, L., and Gheorghiu, D. 2013. Participative teaching for k-12 students with mobile devices and social networks. In *Proceedings of SMART 2013 – Social media in Academia: Research and Teaching*, Bacau, June 7-9, 2013.

Gheorghiu, D., Ştefan, L., and Rusu, A., 2013. E-learning and the process of studying in virtual contexts, pp. 65-93. In Ivanović, M., Jain, L. C. (eds.) *E-learning Paradigms and Applications - Agent-based Approach,* Springer.

Trapp, M., Semmo, A., Pokorski, R., Herrmann, C. D., Döllner, J., Eichhorn, M., and Heinzelmann, M., 2012. Colonia 3D Communication of Virtual 3D Reconstructions in Public Spaces. In *International Journal of Heritage in the Digital Era*, volume 1 number 1.

Veen, W., and Vrakking, B., 2006. *Homo Zappiens. Growing Up in a Digital Age.* Network Continuum Education. London.

Wheat-Stranahan, P., Lippert, D., Van Tuerenhout D., and Phelps, E., 2007. New Ways of Looking at the Past: Archaeological Education at the Houston Museum of Natural Science, pp. 113-126. In Jameson, J., and Baugher, S. (eds.) *Past Meets Present. Archaeologists Partnering with Museum Curators, Teachers and Community Groups*, Springer. New York.

White, M., Liarokapis, F., Mourkoussis, N., Basu, A., Darcy, J., Petridis, P., Sifniotis, M., and Lister, P., 2004. ARCOLite — an XML based system for building and presenting Virtual Museum Exhibitions using Web3D and Augmented Reality, In *Theory and Practice of Computer Graphics 2004 Conference (TP.CG.04)*, Bournemouth.

Winn, W., 1995. The virtual reality roving vehicle project. In *Technological Horizons in Education Journal*, 23(5), pp.70-75.

Winn, W., 2002. What can students learn in artificial environments that they cannot learn in class?, University of Washington, In *First International Symposium*, Open Education Faculty, Anadolu University, Turkey.

www.amire.net

www.europeanvirtualmuseum.net

www.louvre.fr

www.muzeultaranuluiroman.ro

www.timemaps.net/timemap/?page_id=6

Envisioning an Advanced ICT-supported Build-up of Manufacturing Skills for the Factories of the Future

Joao Costa[1], Dimitris Kiritsis[1,2], Poul Kyvsgaard Hansen[3], Manuel Oliveira[4], Loukas Rentzos[5], Hadrien Szigeti[6] and Marco Taisch[7]
[1]*HighSkillz, Kent, U.K.*
[2]*Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
[3]*Allaborg University, Allaborg, Denmark*
[4]*SINTEF, Trondheim Norway*
[5]*University of Patras, Rio Patras, Greece*
[6]*Dassaul Systèmes, Vélizy-Villacoublay Cedex, France*
[7]*Politecnico di Milano, Milano, Italy*
joao.costa@highskillz.com, dimitris.kiritsis@epfl.ch, kyvs@business.aau.dk, manuel.oliveira@sintef.no, rentzos@lms.mech.upatras.gr, hadrien.szigeti@3ds.com, marco.taisch@polimi.it

Keywords: Manufacturing Skills, ICT for Manufacturing, Factories of the Future.

Abstract: In this paper, we present the main concepts and the aim of the ManuSkills FoF project which is to study the use of enhanced ICT-based technologies and training methodologies to facilitate an increase of young talent interest in manufacturing and to support their training of new manufacturing skills. The project will experiment with a wide range of innovative delivery mechanisms such as serious games and teaching factory, supported by the use of social media augmented by gamification and leveraging the distribution channels preferred by young talent. In addition, the project will explore the pedagogical frameworks best suited to the personalization of individual learning needs taking into account the industrial demand. ManuSkills will address all three stages of the young talent pipeline (i.e. children, teenagers, young people), where in the early stages the focus will be to make manufacturing education more attractive to young talent, whilst in the later stages the focus will be to facilitate transformative deep learning of individuals, with reduced time-to-competence.

1 INTRODUCTION

When asking young talents, their understanding of what manufacturing of the future entails could not be further from the truth. In fact, with the shrinking of distances to create the global village coupled with the ever faster pace of advances in technology and process innovation, one can state clearly that manufacturing is in a state of constant innovation. The factory of the industrial revolution, where the shopfloor was unsafe and dirty, is being replaced by factories where lean and TQM have dramatically changed the work environment. A fundamental switch has taken place, where in the beginning the workers carried out repetitious tasks, now the focus is on cognitive interaction with the work environment through ICT. Consequently, higher education has become necessary. In addition to engineering, other competences such as sense making, dealing with complexity and soft skills, are also necessary.

Manufacturing goes digital and this is argued to amount to a third industrial revolution (Economist, 2012). Over the past few decades, manufacturing has evolved from a more labour-intensive set of mechanical processes (traditional manufacturing) to a sophisticated set of information-technology-based processes (advanced manufacturing). The dirty, noisy, unsafe workshops have given place to modern technologies and cleaner work environments. Everything in the factories of the future will be run by smarter software. Digitisation in manufacturing will have a disruptive effect every bit as big as in other industries that have gone digital, such as office equipment, telecoms, photography, music, publishing and films. And the effects will not be confined to large manufacturers; indeed, they will need to watch out because much of what is coming

will empower small and medium-sized firms and individual entrepreneurs. Launching novel products will become easier and cheaper. Communities offering 3D printing and other production services that are a bit like Facebook are already forming online—a new phenomenon which might be called social manufacturing. Things will be made economically in much smaller numbers, more flexibly and with a much lower input of labour, thanks to new materials, completely new processes such as 3D printing, easy-to-use robots and new collaborative manufacturing services available online. The wheel is almost coming full circle, turning away from mass manufacturing and towards much more personalised production. According to the ActionPlanT project vision (ActionPlanT) for future manufacturing – 'Manufacturing 2.0', future ICT-enabled manufacturing will be:

1. **On-demand:** To sustain market share and create employment opportunities, Manufacturing 2.0 should accommodate changing demands from a new customer base and deliver customised products on-demand.
2. **Optimal:** European enterprises need to be able to produce products with superior quality, high security and durability and, at the same time, competitively priced compared to products from emerging markets.
3. **Innovative:** Faster introduction of collective innovation is one of the three key growth factors together with human capital and infrastructures. Innovative thinking, design and manufacturing will lead the way to sovereignty, independence and growth of European manufacturing.
4. **Green:** Manufacturing is responsible for significant energy use and consumption of natural resources. Manufacturing 2.0 needs focused initiatives to reduce energy footprints on shop floors and increase awareness of end-of-life (EoL) product use.
5. **Human-centric:** Manufacturing 2.0 will evolve from being perceived as production centred to human centred with greater emphasis on generating core value for human stakeholders. Future plants should be more accommodating towards the needs of the European workforce and consider them an integral stakeholder.

Moreover, since the manufacturing process, in a broad sense, becomes more and more complex, a big number of complementary roles will be necessary within the Factories of the Future, such as:

- **Virtual enterprise:** Supply chain engineers;
- **Forecasting:** Business analyst and data scientists;
- **Multi technology product:** Industrial engineers;
- **Quality:** Quality managers;
- **More complex manufacturing technologies:** Process engineers and planners;
- **Energy:** Energy managers.

In spite of the constant innovation pushing manufacturing forward, the lack of awareness by mainstream society implies that young talent does not choose a career in manufacturing contributing to a serious shortage of skilled labour. Reaching out to the young talent is the focus of the ManuSkills European project, studying the effectiveness of innovative delivery mechanisms such as serious games and teaching factory, supported by the use of social media augmented by gamification.

The paper is organised as follows: in section 2 the main concepts and project objectives are presented. In section 3 we present the approach we will follow to develop the concepts, implement them in an appropriate platform. In section 4 we present the expected impact to be created by the project actions, achievements and results.

2 CONCEPTS AND PROJECT OBJECTIVES

ManuSkills aims to study how best to reach the young talent using an ICT framework to address all three stages of the young talent pipeline (i.e. children, teenagers, young people), where in the early stages the focus will be to raise awareness about manufacturing to concerned stakeholders, then to make manufacturing education more attractive to young talent, whilst in the later stages the focus will be to facilitate transformative deep learning of individuals, with reduced time-to-competence. The resulting ManuSkills ICT framework is captured in Figure 1 where the triangle of parents, youth and teachers access the framework over their preferred device.

The learning content on the platform is packaged as learning objects provided by both educational institutions and industry, which are delivered using a wide range of mechanisms, such as teaching factory, serious games, webinar, video player and mobile apps. ManuSkills will benefit from existing pedagogical material from within the consortium and study the effectiveness along with the impact on

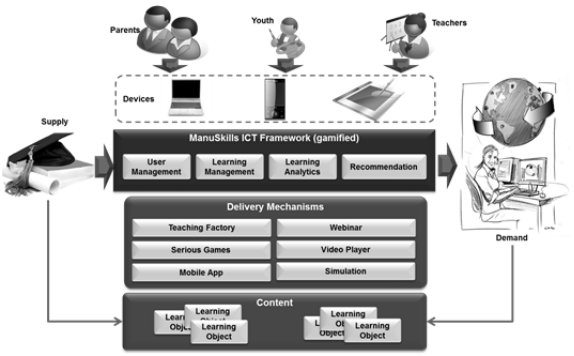


Figure 1: ManuSkills ICT Framework.

both raising the awareness and learning amongst young talent. In fulfilling its aims, ManuSkills will address the following objectives:

- **Objective 1:** Study the barriers and enablers towards raising the awareness of manufacturing and training along the entire young talent pipeline;
- **Objective 2:** Explore how to successfully engage young talent with ICT and increase their interest in manufacturing thus addressing the talent shortage for Europe to sustain its global competitiveness;
- **Objective 3:** Study innovative delivery education mechanisms and training methodologies to facilitate more effective and efficient competence development towards the rapid build-up of existing and emerging manufacturing skills;
- **Objective 4:** Mobilise stakeholders from business, government, academia, international organizations, civil society and nongovernmental organizations to develop action plans and influence policy for improving manufacturing education;
- **Objective 5:** Focus on strengthening manufacturing education at higher education levels and STEM related subjects with focus on engineering in secondary education, addressing the three target stakeholder groups: children, parents and educators.
- **Objective 6:** Achieve early impact based on the ManuSkills proof-of-concept platform and other project results, and facilitate wider impact through aggressive dissemination and networking activities supported by a **sustainable exploitation case.**

3 APPROACH

A framework capable of capturing the existing and future manufacturing skills, could be the basis of a, currently lacking, reference to professional needs, related to advanced manufacturing. These skills are essential to manufacturing companies in order to improve their innovativeness and competitiveness. Skilled young talent will strengthen the capacity to innovate and will give new incentive to industry in order to hinder the outflow of manufacturing activities to countries outside Europe. The synthesis of such a reference framework of manufacturing skills will be updateable in order to allow for new skills to be added that may result from new developments in advanced manufacturing. This framework will be used as a reference for manufacturing education and will enable capturing the main manufacturing skills, required by European companies for the promotion of innovation and creativity. The development of this reference framework will be based on a taxonomy of manufacturing skills with reference to existing relevant studies. This taxonomy will identify the main classes of manufacturing skills needed to compete in the evolving manufacturing ecosystem. It will comprise of mainly three categories, namely processes, equipment and systems. These main classes include manufacturing skills related to knowledge and optimization, control and human-machine interfaces, logistics and information flow, creativity and innovation.

On the basis of this framework, the project will define a vision on new approaches for awareness creation and training for the rapid build-up of manufacturing skills, using ICT enabling technologies. This vision will be reflected on a roadmap that will identify and give insight on the training technologies and approaches that will attract the young talent in manufacturing education and create engineers with manufacturing skills required by the factories of the future. The roadmap will be coordinated with both European SME ICT providers and users, for the development of future manufacturing skills, related to advanced manufacturing. This roadmap will have to drive the knowledge delivery mechanisms and the ICT technologies needed for assimilating the identified manufacturing skills, in the context of the future factories. In addition, the project will undertake an aggressive social awareness of the roadmap by European society, resorting to both traditional media (e.g. broadcast television, internet, press, etc.) and social media strategies with viral impact (i.e. social

networks, gamified apps, etc.). Existing ICT tools that can contribute to the rapid build-up of manufacturing skills will have to be enhanced and extended, while the need for new tools will be identified. The aim will be to boost competitiveness and innovation in European manufacturing that would be requiring a workforce with advanced manufacturing skills.

3.1 Additional Steps

Some of the required steps to bring about desired impact have been already mentioned in the previous paragraphs. Additional steps that ManuSkills will take are summarized as follows:

- ManuSkills project will involve relevant stakeholders as part of the development and deployment process, i.e. there will be meetings, seminars and other events with representatives from educational and research authorities and associations as well as for the relevant interest groups (e.g. primary/secondary school teachers, parents representatives, researchers, university professors) within the STEM education.
- Building on existing professional networks of teachers and teacher associations will provide opportunities to disseminate ManuSkills results and furthermore foster cooperation between schools, collaborative reflection, exchange of ideas and experiences and quality development.
- There will be constant communication of results and participation in discussions through relevant channels (such as newspaper, television, journals, conferences, web blogs etc.) through the whole duration of the project. The main results emerging from ManuSkills project are very likely to get a wide audience and a direct impact on debates in national and international contexts.
- Most of the project partners have already established networks with high schools, universities and industry in their countries and have been purposely chosen to participate and contribute to ManuSkills in order to build the community of stakeholders and reach the desired impact through established channels. The stakeholders within established networks all recognize the importance of STEM in a knowledge based society and the urgent need to enhance young people's interest towards STEM.

3.2 European Added Value

ManuSkills needs to take into the consideration the large European context as the TEL solutions to be implemented in schools, which are applicable within one context (i.e. within one country), could not simply be transferrable to other contexts (i.e. other countries). ManuSkills project is also following the initiative "*Science Education NOW*" (EU, 2007) which recommends that the articulation between isolated national activities in STEM education and those funded at the European level must be improved.

Several barriers can emerge in the deployment process of the ManuSkills ICT Framework (such as infrastructural barriers, accessibility barriers, legislative and education policy obstacles, lack of means to ensure long-term sustainability etc.) and in the European context; all these barriers strongly differ from country to country. ManuSkills will do the analysis of the present state of the different educational systems and corresponding institutional practices in the partners' countries to identify country specific drivers and barriers (e.g. related to curriculum and pedagogy, infrastructure, capacity-building, language and content, financing etc.) to ensure the wide implementation of the ManuSkills results in different European contexts.

4 EXPECTED IMPACT

The expected higher level impacts of the ManuSkills project are the following:

- Penetrate new application areas (e.g. high customisation, end-of-life product engineering and manufacturing), close to the market and opening new markets;
- Strengthen supply-side SMEs by enabling them to supply manufacturers with new equipment and components for improved manufacturing operations;
- Leveraging innovation capacity and competitiveness of European producers of laser manufacturing equipment and their suppliers, in particular SMEs, and of the users of such equipment;

ManuSkills will contribute only indirectly to the above impacts through the preparation of new generations to become eager for developing skills in advanced manufacturing and contribute in the increased competitiveness of European manufacturing industry.

We think that the main strategic impact generated by ManuSkills will be addressing the shortage of talented skills in manufacturing. The

World Economic Forum report "*The Future of Manufacturing. Opportunities to drive economic growth"* (WEF, 2012) is stated:

"An estimated 10 million jobs globally with manufacturing organizations cannot be filled today due to a growing skills gap. Despite the high unemployment rate in many developed economies, companies are struggling to fill manufacturing jobs with the right talent. Emerging nations cannot fuel their growth without more talent. Access to talent will become more important and more competitive. The skills gap that exists today will not close in the near future, which means companies and countries that can attract, develop and retain the highest skilled talent – from scientists, researchers and engineers to technicians and skilled production workers – will come out on top. In the race to future prosperity, nothing will matter more than talent."

We believe that ManuSkills main impact will be to contribute in reducing this gap, by increasing in the young talented people the awareness of a potential interesting professional career in the manufacturing sectors.

Anyhow more detailed direct impacts created by ManuSkills will be generated by the produced results in **medium-long term** and short, term as follows. Being medium-long term the quantification of the impact is vepry difficult to be done.

- To increase the competitiveness of the European companies by providing them the talented skilled workers;
- Innovation is more and more dependent on talented scientists and technicians. ManuSkills will help have an impact on the *Europe 2020 Flagship Initiative* on *Innovation Union*;
- To raise the awareness of publishers and learning content providers towards the growing demand of adapted manufacturing education taking into accounts the rapid evolution of technology and the inspirations of the young people. This will increase the quantity and the quality of the education content available in internet and easy to be reached;
- Systematic approach for continuous reinforcement of manufacturing by emerging technologies through developing strategies for providing appropriate training to acquire the related skills;
- Putting the humans at the centre of interest in manufacturing to remain competitive and innovative through institutionalisation of continuous and lifelong learning of the employees to improve their skills and increase their capabilities.

5 CONCLUSIONS

This envisaged evolution brings new needs for skills development. The adaptation of the educational content and its delivery mechanisms to the new requirements of knowledge-based manufacturing, the provision of integrated engineering competencies, including a variety of soft skills, the promotion of the innovation and entrepreneurship spirits, as well as an interdisciplinary thinking, reflecting the increasing integration of different areas of knowledge in manufacturing, are considered as major priorities.

The ActionPlanT project has delivered a validated set of recommendations and guidelines for an effective development of ICT for manufacturing skills. On top of that, the ManuSkills project will deliver an ICT framework and a roadmap for manufacturing awareness and skills development

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 609147.

The work in the ManuSkills project is a common effort among all its contributing partners: Aalborg University, EPFL, Dassault Systèmes, HighSkillz, POLIMI, SINTEF and University of Patras.

REFERENCES

ActionPlanT (2011), ICT FOR MANUFACTURING: The ActionPlanT Vision for Manufacturing 2.0.

Economist (2012), A third industrial revolution, Special report: Manufacturing and innovation, April 2012, http://www.economist.com/node/21552901.

EU (2007). *Science Education NOW! – Renewed Pedagogy for the Future of Europe.* Brussels: European Commission, High Level Group on Science Education.

WEF (2012), The Future of Manufacturing. Opportunities to drive economic growth, World Economic Forum. April 2012.

Development and Practical Application of a Relationship Diagram-creation Tool Centering on Automatic Creation Functionality

Norio Ishii[1], Yuri Suzuki[2] and Saori Sakuma[3]
[1]*Aichi Kiwami College of Nursing, 5-4-1 Jogan-dori, Ichinomiya, Aichi, Japan*
[2]*College of Engineering, Chubu University, 1200 Matsumoto-cho, Kasugai, Aichi, Japan*
[3]*School of Nursing, Seirei Christopher University, 3453 Mikatahara-cho, Kita-ku, Hamamatsu, Shizuoka, Japan*
n.ishii.t@aichi-kiwami.ac.jp, yuris@cs.chubu.ac.jp, saori-s@seirei.ac.jp

Keywords: Nursing Education, Learning Support Tool, Relationship Diagram, Automatic Creation.

Abstract: In nursing education, learning the effects of relationship diagrams drawn by students is useful in understanding patients and the entire nursing process. However, drawing these diagrams is time consuming, and the diagram creation process is strenuous for students. In this study, we have developed a relationship-diagram creation tool that further reduces the difficulty of node organization. In particular, we have developed a tool that implements an automatic drawing function in which the tool, rather than the student, organizes information in the diagram. We conducted lectures on methods to use the tool at a nursing college as well as a questionnaire survey to evaluate its efficacy. On the basis of these results, we confirmed the improvements in the tool's usability and efficacy.

1 INTRODUCTION

The creation of relationship diagrams is important in the nursing process. Relationship diagrams are "concept maps" or "concept networks" that logically use arrows to depict the causes of illness, organic changes, functional changes, symptoms, and decreasing lifestyle activity to understand a patient. Relationship diagrams are used in basic nursing education to organize and integrate related information and present an overall image of a patient. They are important in determining the course the nursing process takes (Akinsanya and Williams, 2004; Hsu and Hsieh, 2005).

Earlier research has reported the use of relationship diagrams in nursing education to reinforce logical thinking and understanding (Sugisaki and Ogawa, 2006; Yakushijin et al., 2006). This study has noted that consolidating information improves learning and critical thinking abilities. However, it has been pointed out that the placement of nodes and lines in relationship diagrams make them difficult to draw and revise (Sugisaki and Ogawa, 2006). This suggests that, for students, the creation of relationship diagrams is a complex and difficult activity that requires an understanding of the patient, an ability to follow the nursing process, and critical thinking capabilities.

As the earlier study suggested, creating relationship diagrams has merits, such as (1) promoting understanding of the patient as a whole and (2) improving logical and critical thinking capabilities. The current study focuses on (1) and aims to construct a learning environment for beginners to support the understanding of the patient as a whole.

Earlier studies tested a number of methods to increase the effectiveness of relationship diagrams as a learning tool, as described in (1), by avoiding confusing the students with detailed rules, including giving students rules for drawing diagrams that distinguish between actual and potential conditions and depicting different elements of information with different colors (All and Havens, 1997; Schuster, 2002; Toyoshima et al., 2005). These practical studies can be seen as a research approach that supports students by proposing effective rules for drawing relationship diagrams, which have not had clear and established drawing rules.

In contrast, the authors have performed research based on the development of a learning support tool for creating relationship diagrams and constructing a

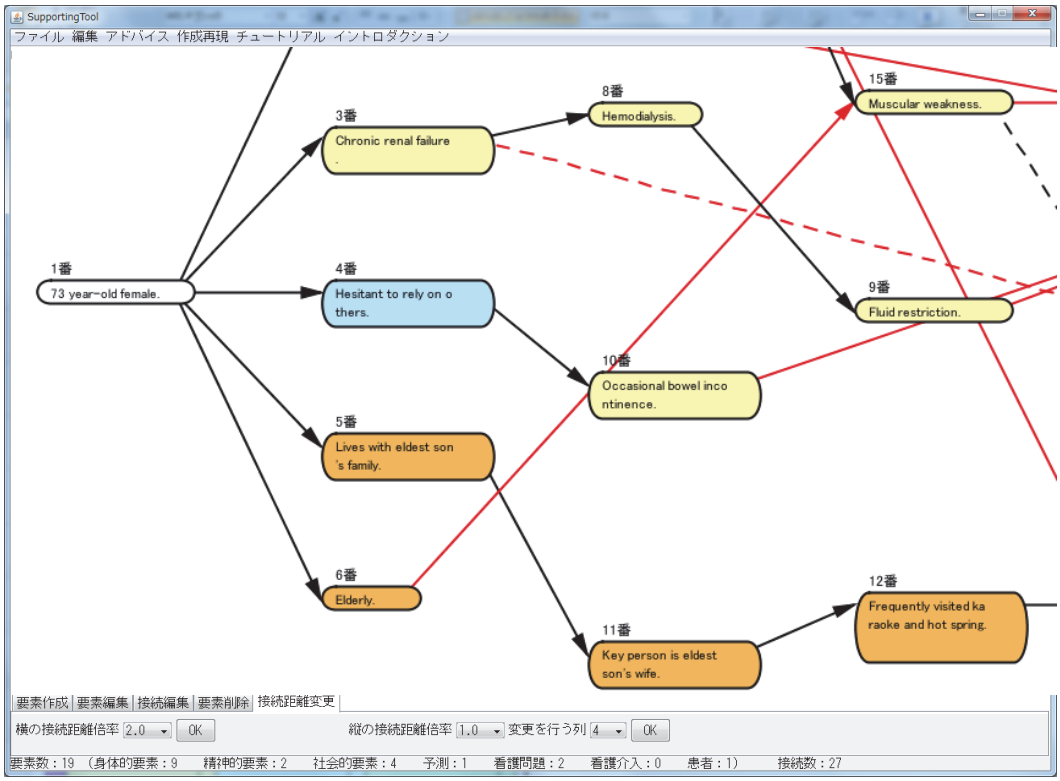


Figure 1: Relationship-diagram creation tool.

learning environment that reduces the previously mentioned student difficulties. In previous research, we developed a computer-based tool for simple creation of relationship diagrams using Microsoft Excel VBA and introduced this program into the classroom in an attempt to reduce the burden of diagram creation (Ishii and Sakuma, 2011). Using this tool allowed students to easily create and organize relationship diagram nodes and links. In practical use, the tool received positive evaluations from students in terms of creating and editing diagrams. However, no difference was seen in the amount of time required to create a diagram with the tool and the time to create a diagram by hand. We inferred that this was because the time spent by students on editing tool-generated diagrams was greater than that spent on hand-drawn diagrams since the tool allowed students to easily rearrange nodes.

In this study, we have developed a relationship diagram-creation tool that further reduces the burden of node organization. Specifically, we developed a tool that implements an automatic drawing function in which the tool rather than the student organizes information in a diagram. We then deployed this program in a classroom. In addition, we made improvements to the tool based on student evaluations, and we investigated the effects of the improvements by introducing the program in the classroom once again during the following academic year.

2 A RELATIONSHIP DIAGRAM-CREATION TOOL

2.1 Development Environment

We developed the tool in the Java development language under the Windows operating system. During development, we used three libraries: Processing, which is specialized for graphics functions; Apache POI, which reads and writes files in formats for Microsoft applications such as Excel and Word; and JFreeChart, which allows creation graphs from Java applications.

2.2 Tool Functions

In this study, we developed a relationship diagram-creation tool with three functions: an automatic drawing function, a creation process replay function, and an evaluation support function (Figure 1).

2.2.1 Automatic Drawing Function

The tool automatically creates diagrams based on information entered by the student. We call this the automatic drawing function. The diagrams are composed of nodes and links, and nodes are hierarchically organized from left to right. Relationship diagrams are generally composed in a radial shape; however, based on interviews with nursing instructors, we have adopted a hierarchical composition, which was assessed as easy to read.

To create a node, the user selects the "create element" tab from the lower part of the tool. After inputting the target information for the new element, the user selects element properties and connecting element numbers (a node number pointing to a higher level in the hierarchy) from the menu. This automatically creates a new node in a color based on its properties and places the node in the diagram. The node is also automatically assigned a number in order of creation. There are six element properties: (1) physical attributes (height, weight, ailments, etc.), (2) social background elements (family composition, household environment, etc.), (3) psychological state elements (mental state, mental issues, etc.), (4) expectations (future expectations regarding the patient), (5) nursing obstacles (critical issues concerning nursing), and (6) nursing intervention (nursing activity for the subject patient).

To edit node information, the user selects the "edit element" tab, and then selects the number of the node they wish to edit from the pulldown menu. Current information is shown in the content field of the element. Current information can be edited. When the OK button is pressed, these changes are reflected in the diagram. Nodes can be deleted by selecting the "delete node" tab and selecting the number of the node to be deleted from the pulldown menu. However, to preserve links, only nodes with no lower level nodes can be deleted.

Links between nodes are automatically generated, and normal links are indicated by solid arrows. Prediction links are indicated by dotted arrows. To add a link between nodes, the user chooses the "edit connections" tab and enters edit mode to select the number of nodes to be connected from the pulldown menu. The added link is shown with a red arrow. To delete a node, after selecting deletion mode, the user must then choose the number of the node to be deleted in a similar fashion.

Finally, as nodes and links are created automatically, the student cannot change their placement; however, by selecting the "change connection distance" tab, the users can adjust the spacing between nodes.

2.2.2 Creation Process Replay Function

The tool automatically logs the creation process of the relationship diagram and can recreate this process step by step after creation. This is called the creation process replay function. The creation process replay includes the deletion as well as addition of elements and connections. This function allows students and instructors to confirm a student's creation process.

2.2.3 Evaluation Support Function

After the creation of a relationship diagram, this tool can display an evaluation sheet that shows the number of links and nodes of each type in a table as well as the proportion of nodes of each type as a graph. This is called the evaluation support function. Evaluation sheets are output in a Microsoft Excel-compatible format.

3 INTRODUCING THE TOOL IN THE CLASSROOM

3.1 Course and Object of Study

In this study, we introduced the developed relationship diagram-creation tool into a classroom. The target course was "The Nursing Process," a first year open lecture at a short-term nursing school. One of the course's fifteen total classes focuses on the creation of relationship diagrams. There were 88 first-year short-term nursing college (school of nursing) students in the class.

3.2 Class Design

The class content was composed of three steps: introduction of the tool, creation of relationship diagrams, and evaluation and improvement of diagrams.

In the first step, students received a basic explanation of the fundamentals of the tool. Specifically, students received an explanation regarding the creation, revision, and deletion of nodes as well as the addition of nodes and adjustment of spacing between nodes. The procedure for saving files was also explained.

In the second step, students were divided into groups of two to three and created one full image

Table 1: Evaluation items.

No.	Usability	Efficiency
1	I think that I would like to use this tool frequently.	It didn't take long.
2	I found the tool unnecessarily complex.	It was possible to draw diagrams easily.
3	I thought the tool was easy to use.	It was possible to draw diagrams neatly.
4	I think that I would need the support of a technical person to be able to use this tool.	It was possible to draw diagrams in detail.
5	I found the various functions in this tool were well integrated.	It was easy to erase nodes or links.
6	I thought there was too much inconsistency in this tool.	It was easy to revise diagrams.
7	I would imagine that most people would learn to use this tool very quickly.	It was easy to rearrange nodes or links.
8	I found the tool very cumbersome to use.	It was useful to classify information by kinds of color or line.
9	I felt very confident using the tool.	It was easy to organize information.
10	I needed to learn a lot of things before I could get going with this tool.	It was easy to consider relationship.

relationship diagram based on a given case (symptoms associated with stroke, physical changes associated with aging, and the effects of these symptoms on daily life). Because the students are beginners at creating diagrams, we distributed a data file in which one portion related to the illness was created in advance. Students completed a relationship diagram by adding information to the distributed diagram.

In the third step, students used the evaluation sheet provided by the tool to assess the created relationship diagram. Students then added supplementary information and revised incorrect information. Finally, students reviewed the creation process using the creation process replay function.

After the diagram creation process finished, students assessed the tool following the same process used in earlier research (Ishii and Sakuma, 2011). More specifically, a system usability scale (Brooke, 1996) (10 items, 5-rank evaluation) was used to assess usability. Evaluation items based on points that the students considered difficult in earlier research were also used (10 items, 4-rank evaluation) to assess efficiency. The evaluation items are shown in Table 1.

3.3 Student Evaluation of the Tool

In our evaluation of the tool, we first compared evaluations received from students in this study with those obtained in earlier research (Ishii and Sakuma, 2011) regarding the previous relationship diagram tool. As explained in Section 1, the tool in this research varied from the previously developed tool as follows: (1) nodes were automatically created by entering information but students manually organized the diagram; (2) similar to nodes, links were also manually organized by the students; and (3) the creation process replay functionality and evaluation support functionality were not implemented. In addition, creation of the relationship diagram was performed over two classes in previous studies, whereas in this study, the relationship diagram was created in a single class period.

3.3.1 Usability Evaluation

Figure 2 shows usability evaluations for the previous and proposed tools. The system usability scale utilized in this study included inverted evaluation items. In Figure 2, positive evaluations are represented with higher values.

Figure 2 shows that the evaluations were higher for the proposed tool than the previous tool for all 10 evaluation items. A non-corresponding t-test performed for the average difference in total usability points between the previous and proposed tools confirmed that the proposed tool obtained a significantly higher average point total (33.9) than the previous tool's average (31.1) ($t(146)=3.572$, $p<.01$). However, the average point values were less than 3.0 for items 4, 9, and 10 in the evaluations of the proposed tool.

3.3.2 Efficiency Evaluation

Figure 3 shows efficiency evaluation results for the proposed and previous tools. A non-corresponding t-test performed for the average difference for items 1 to 10 confirmed that the proposed tool obtained a significantly lower average than the previous tool's

average for item 7 (t(146)=2.377, p<.05). While the proposed tool received higher point totals in most evaluation categories, we did not find a significant difference.

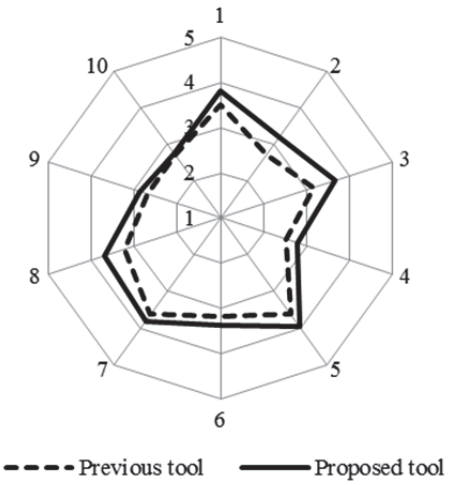


Figure 2: Result of usability evaluations using the previous and proposed tools.

3.4 Evaluations of Student Created Relationship Diagrams

To show the effect of the proposed tool on learning, we evaluated the relationship diagrams created by students and conducted student evaluations of the tool. This study established evaluation items for (1) the number of nodes for each type and (2) the number of links (excluding the crossing links mentioned below), which express the amount of information in the diagram, and (3) the number of crossing links (links that cross node levels; we observed activity when links were added to previously created nodes), which are used as an evaluation item in representative earlier research (Novak and Gowin, 1984). The subjects of analysis were relationship diagrams (14 from the previous tool, 30 from the proposed tool) created by consenting student groups.

Table 2 shows the results of analyzing the relationship diagrams created in experiments using the previous and proposed tools. Each item is represented by a number; however, this excludes data included in data files provided at the beginning of diagram creation. A non-corresponding t-test performed for the average difference for each evaluation item confirmed that diagrams created by the proposed tool had significantly more expectation nodes (t(42)=2.330, p<.05) and crossing links.

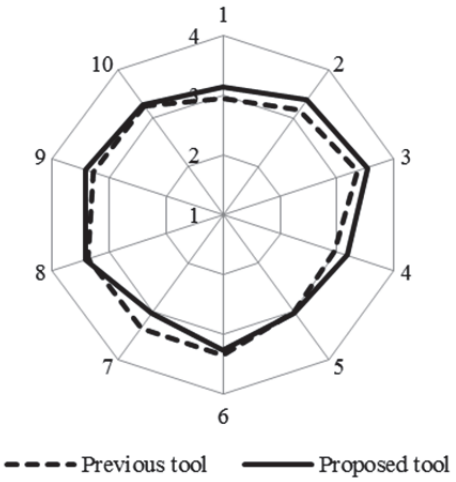


Figure 3: Result of efficiency evaluations using the previous and proposed tools.

Table 2: Results of analyzing the relationship diagrams using the previous and proposed tools.

Evaluation items	Previous tool	Proposed tool
Physical attributes	4.7	5.2
Social background elements	1.6	0.3
Psychological state elements	2.9	3.7
Expectations	1.1	2.2
Nursing obstacles	1.7	1.3
Nursing intervention	2.8	3.0
Links	15.7	14.8
Crossing links	2.9	6.8

(t(42)=3.865, p<.01) compared with those created by the previous tool. We also found that the relationship diagrams created with the previous tool had a significantly larger number of nodes regarding social background elements (t(42)=4.505, p<.01).

3.5 Discussion

Usability evaluations of the tool confirmed that the proposed tool received higher overall evaluations than the previously developed tool. This suggests that the control methods and interface developed in this study were easy for the students to understand.

No significant difference between the previous and proposed tool was seen in evaluation of tool efficiency. Low evaluation of the proposed tool's ability to change the organization was due to the specifications of the proposed tool; students could not reorganize the diagram. As no difference was seen in other items, we hypothesize that this specification did not have a large overall effect on relationship diagram creation.

Analysis of the relationship diagrams created by students confirmed that the number of prediction nodes and crossing links increased with the proposed tool. This suggests that students actively investigated the possibility of future occurrences based on the provided information as well as the relationships among the information. Diagram creation required two class periods with the previous tool, whereas the proposed tool used only one. This indicates the efficiency of the proposed tool.

Future improvements could include strengthening support functions for promoting link creation and diagram nodes. In particular, the results showed that few social background elements were created by the proposed tool. In future, we aim to add additional functions to promote entry of such elements.

4 IMPROVING THE RELATIONSHIP DIAGRAM-CREATION TOOL

4.1 Adding Functions to the Tool

To increase the efficiency of the relationship diagram-creation tool introduced to the classroom in Section 3, we have implemented the following additional functions: an initial information entry function and a nursing problem extraction function.

4.1.1 Initial Information Entry Function

The initial information entry function supports activity at the beginning of the relationship diagram creation process. Students enter basic information regarding the patient using a template displayed in list format (age, sex, family composition, occupation, medical history, and disease name). This data is saved as an initial information file. By importing the initial information file through the tool menu, a relationship diagram based on all the information is displayed.

This function is expected to promote the creation of social background elements and physical attributes related to the target patient. We also expect this function to increase productivity of student data entry during the initial creation of the relationship diagram, reducing both data omissions and time required to create a diagram.

4.1.2 Nursing Problem Extraction Function

The nursing problem extraction function supports evaluations after creation of the relationship diagram. As stated in Section 2.2.3, we implemented an evaluation support function that provided quantitative information related to the relationship diagram. However, evaluating the content of the created relationship diagram was difficult with this function. Thus, we implemented the nursing issue extraction function to assess the content of the relationship map effectively. By selecting this function, links are extracted from the initial information nodes to the nursing issue (crucial issues in nursing) nodes and displayed in another window (for example, “76-year-old woman > stroke > incontinence > damp skin > bedsores).

This function is expected to make it easy for students to conduct an evaluation by reviewing their thought process through a large quantity of information, which can promote understanding of the subject.

4.2 Introducing the Improved Tool into the Classroom

We introduced the improved tool in a class to assess the effects of the improvements. As with the experiment described in Section 3.1, we used one class period of “The Nursing Process,” a first-year open lecture at a short-term nursing university. There were 82 first-year short-term nursing college (school of nursing) students in the class.

The class design was approximately the same as the previous class, with the following changes. (1) We added an explanation of the initial information entry function at the first step. (2) In the second step, students created a relationship diagram using the initial information entry function (students were given a printed sheet of information rather than a data file). (3) In the third step, the relationship diagrams were assessed using the nursing issue extraction function in addition to the evaluation sheet and creation process replay function. After the class, tool evaluation was performed using the same questionnaire. 66 students were targets of this analysis.

4.3 Student Evaluations of the Improved Tool

In Section 3, we assessed the tool through comparison with the previously developed tool. Here we compare student evaluations of the proposed tool before and after implementing the improvements described in Sections 4.1.1 and 4.1.2.

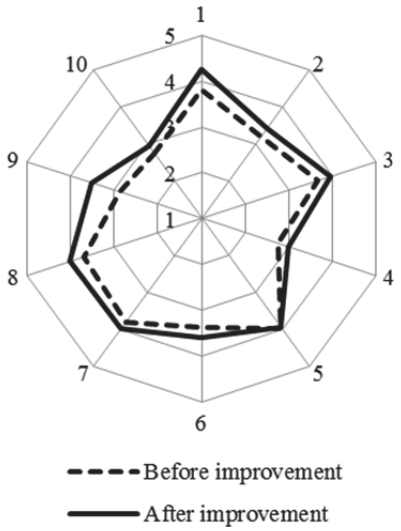


Figure 4: Result of usability evaluations of the tool before and after improvement.

4.4 Student Evaluations of the Improved Tool

In Section 3, we assessed the tool through comparison with the previously developed tool. Here we compare student evaluations of the proposed tool before and after implementing the improvements described in Sections 4.1.1 and 4.1.2.

4.4.1 Usability Evaluation

Figure 4 shows the usability evaluations of the tool as points before and after improvement. Figure 4 shows that the evaluations of the tool after improvement were higher in every item except 5. A non-corresponding t-test performed for the average difference in total points of usability for the tool before (33.9) and after improvement (36.6) showed that the improved tool received significantly higher average point totals ($t(135)=2.855$, $p<.01$). Items 4, 9, and 10, which received evaluations with average point scores below 3.0 prior to improvement, rose above 3.0 after the improvements were implemented.

4.4.2 Efficiency Evaluation

Figure 5 shows efficiency evaluation results of the tool before and after improvement. A non-corresponding t-test performed for the average difference in points for each evaluation item confirmed that the improved tool received significantly higher evaluations in items 1 ($t(144)=2.365$, $p<.05$), 2 ($t(144)=3.706$, $p<.01$), and10 ($t(133)=2.882$, $p<.01$). No evaluations were significantly lower after improving the tool.

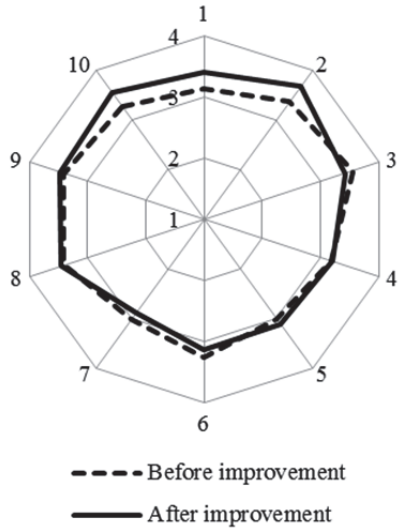


Figure 5: Result of efficiency evaluations of the tool before and after improvement.

4.5 Evaluation of the Relationship Diagrams Created by Students

As with the relationship diagrams created with the proposed and previous tools, we assessed the diagrams created before and after improvement of the proposed tool. The subjects of analysis were relationship diagrams (30 with the unimproved tool, 20 with the improved tool) created by consenting student groups.

Table 3 shows the results of analyzing the relationship diagrams created using the tool before and after improvement. As in Table 2, the numerical values exclude information included in the data file. A non-corresponding t-test performed for the average difference in points for each item confirmed that diagrams created with the improved tool contained significantly higher nodes containing physical attributes ($t(48)=3.751$, $p<.01$) and social background elements ($t(48)=2.317$, $p<.05$). In contrast, we found that psychological state elements were significantly more numerous in the relationship diagrams created with the tool prior to improvement ($t(48)=2.911$, $p<.01$).

4.6 Discussion

Usability evaluations confirmed that the improved tool received higher overall evaluations than the tool prior to improvement. This suggests that the additional functions improved usability even further

when creating new relationship diagrams and assessing the created diagrams.

Table 3: Results of analyzing the relationship diagrams using the tool before and after improvement.

Evaluation items	Before improvement	After improvement
Physical attributes	5.2	8.8
Social background elements	0.3	0.9
Psychological state elements	3.7	2.3
Expectations	2.2	1.7
Nursing obstacles	1.3	0.7
Nursing intervention	3.0	3.5
Links	14.8	15.7
Crossing links	6.8	7.6

In addition, efficiency evaluations of the tool showed improvement in three evaluation items; in particular, 1, which considered creating diagrams, improved. The reason for this is thought to be that by adding the initial information entry function, it was possible to reduce the time required for students to enter information. In addition, 10, which considered the ease of thinking about relationships, also improved. This suggests that the nursing issue extraction function was effective in making relationships easier to grasp.

Analysis of the relationship diagrams created by the students confirmed that physical attributes and social background elements increased in the improved tool. This is thought to be because the improved tool reflects the results of continuing creation while the student focuses on social background elements, such as occupation and family composition, and physical attributes, such as medical history, through the initial information entry function implemented in the tool. We hypothesize that the reduction in psychological state elements was due to the absence of psychological state elements in the initial information.

In this study, during creation of the initial information we selected the basic information that we wanted students to enter in the relationship diagram. If the selected information had a strong effect on the content of the relationship diagram, we considered that it should include information that would serve to develop opportunities for further consideration; for example, problems during hospitalization and psychological state elements, such as what the patient does not want to do. In future, we aim to investigate this point through practical experiments.

5 CONCLUSIONS

The relationship diagram-creation tool developed in this study received more positive evaluation than tools developed in earlier research. We have also examined the effects of further improving the tool.

While this study conducted an analysis of the relationship diagrams created by the students, future research is aimed at performing detailed analysis of the creation process and experimentally investigating the effect of each function implemented to facilitate addition of knowledge obtained in the current study. Based on the results of this analysis, we hope to improve the tool further and create a more effective learning environment.

REFERENCES

Akinsanya, C., Williams, M., 2004. Concept mapping for meaningful learning. *Nurse Education Today*, 13(1), 41-46.

All, A., Havens, R., 1997. Cognitive/Concept mapping: A teaching strategy in nursing. *Journal of Advanced Nursing*, 25(6), 1210-1219.

Brooke, J., 1996. SUS: A “quick and dirty” usability scale. In Jordan, P. W. et al., (eds.), *Usability Evaluation in Industry*. Taylor and Francis.

Hsu, L., Hsieh, S., 2005. Concept maps as assessment tool in a nursing course. *Journal of Professional Nursing*, 21(3), 141-149.

Ishii, N., Sakuma, S, 2011. Supporting the creation of sequence of events in nursing education. *International Journal of Education and Information Technologies*, 5(2), 199-206.

Novak, J. D., Gowin, D. B.,1984. *Learning How to Learn*. Cambridge University Press.

Schuster, P. M., 2002. *Concept Mapping: A Critical Thinking Approach to Care Planning*. F.A. Davis.

Sugisaki, H., Ogawa, I., 2006. Learning effects of use of sequence of events in adult nursing practicum (Acute period). *Bulletin of Nara Medical University School of Nursing*, 2, 1-6.

Toyoshima, Y., Itou, F., Hagi, Y., Nishibori, Y., Kazaoka, T., Kishita, S., Itou, S., 2005. Evaluation of student learning about the nursing process using written simulation in adult nursing course (Part 3): Analysis of student self-evaluation in learning with sequence of events. *Bulletin of Department of Nursing Seirei Christopher College*, 13, 81-90.

Yakushijin, Y., Murakami, E., Nakamura, K., 2006. Learning effectiveness of using concept maps in pediatric clinical nursing practice. *Journal of Japanese Society of Child Health Nursing*, 15(2), 8-14.

AUTHOR INDEX

Adachi, T. 252
Alba-Elías, F. 34
Alenyà, G. 213
Aoki, K. 139
Balksjö, T. 336
Banerjee, S. 363
Bele, J. 125
Bensassi, M. 81
Bergen, H. 5
Blesa, M. 159
Blewitt, W. 95
Boas, I. 205
Boork, M. 336
Boytchev, P. 264
Bracey, G. 192
Brante, E. 309
Brito, T. 110
Calderón, A. 117
Campagni, R. 26
Campos, G. 317
Chen, L. 298
Clemmensen, L. 280
Costa, J. 389
Cruz, D. 205
Daskolia, M. 342
Dey, P. 363
Dhanasekar, M. 275
Dias, M. 197
Dimc, M. 125
Doxas, I. 301
Dubois, M. 5
Duch, A. 159
Engström, S. 336
Ersbøll, B. 280
Essalmi, F. 147
Feenstra, W. 270
Feng, B. 298
Flores, A. 246
Fonte, D. 205
Funaoi, H. 258
Gabarró, J. 159
Gabi, C. 197
Gançarski, A. 205
Garvin-Doxas, K. 301
Gheorghiu, D. 379
Giordano, D. 42
González-Marcos, A. 34
Granados, A. 103
Grosch, P. 213
Hansen, P. 389
Harada, K. 139
Hatziapostolou, T. 57
Henriques, P. 205
Henz, M. 14
Hernández, S. 213
Hettiarachchi, E. 184
Hilliard, A. 288
Huertas, M. 184
Inagaki, S. 252, 258
Ishii, N. 394
Izuishi, K. 252
Jang, E. 226
Jemec, A. 125
Jemni, M. 147
Kang, E. 329
Kang, S. 226
Khenissi, M. 147
Kim, M. 226
King, S. 231
Kinshuk, ... 147
Kintsch, W. 301
Kiritsis, D. 389
Klymkowsky, M. 301
Kogan, J. 221
Krampen, G. 353
Krishnan, G. 67
Kusunoki, F. 252
Kynigos, C. 342
Laroussi, M. 81
Lazarov, B. 153
Lehmann, P. 235
Leichner, N. 353
Letón, E. 292
Liu, J. 270
Liu, L. 231
Liu, N. 298
Liu, X. 275
Locke, S. 192
Lopes, R. 131
Magrane, K. 270

AUTHOR INDEX (CONT.)

Maiorana, F. 42, 49
Mandal, K. 363
Mani, A. 176
Marlette, S. 192
Martín-Carrasco, F. 103
Matei, S. 371
Mayer, A. 353
McGee, K. 14
Mediero, L. 103
Merlini, D. 26
Mesquita, C. 131
Mizoguchi, H. 252
Molanes-López, E. 292
Mor, E. 184
Morgan, G. 95
Moura, J. 110
Muratsu, K. 258
Murayama, S. 139
Nakayama, T. 252
Navaridas-Nalda, F. 34
Nikolova, N. 264
Noh, E. 226
Nyström, M. 309
Ogitsu, T. 252
Olander, M. 309
Oliveira, M. 389
Oliveira, N. 205
Olsen, R. 336
Ordieres-Meré, J. 34
Otero, A. 270
Ovaska, E. 89
Paes, J. 110
Paraskakis, I. 57
Peacock, K. 231
Perkusich, A. 197
Peter, J. 353
Petit, J. 176
Pirhonen, M. 323
Poonnawat, W. 235
Powell, N. 167
Putter, L. 5
Queinnec, C. 241
Raibulet, C. 89
Rann, J. 167
Rentzos, L. 389
Rivero, J. 213
Rodrigues, E. 197
Rodríguez, M. 184
Roque, G. 317
Roura, S. 176
Rozman, D. 125
Rughinis, C. 371
Rughinis, R. 371
Ruiz, M. 117
Rull, A. 213
Sakuma, S. 394
Santillan, D. 103
Sengupta, P. 67
Seong, T. 226
Serna, M. 159
Shabanova, M. 153
Singhal, R. 14
Sliusarenko, T. 280
Souza, V. 197
Ştefan, L. 379
Stefanov, K. 264
Stefanova, E. 264
Stembridge, D. 167
Sung, K. 226
Suzuki, Y. 394
Szigeti, H. 389
Taconis, R. 5
Taisch, M. 389
Takemura, H. 252
Ushaw, G. 95
Venkataramani, D. 176
Verri, M. 26
Wright, M. 67
Yamaguchi, E. 258
Yang, S. 298
Yoon, I. 329
Yoshida, R. 252
Yu, H. 167
Zong, N. 275

Proceedings of CSEDU 2014 - Volume 2
6th International Conference on Computer Supported Education
ISBN: 978-989-758-021-5 | www.csedu.org

Copyright 2014 SCITEPRESS
Science and Technology Publications
All Rights Reserved

INSTICC IS MEMBER OF:

LOGISTICS PARTNER:

PROCEEDINGS WILL BE SUBMITTED FOR INDEXATION BY: