Educational Technology & Society
An International Journal

Aims and Scope

Educational Technology & Society is a quarterly journal published in January, April, July and October. Educational Technology & Society seeks academic articles on the issues affecting the developers of educational systems and educators who implement and manage such systems. The articles should discuss the perspectives of both communities and their relation to each other:

- Educators aim to use technology to enhance individual learning as well as to achieve widespread education and expect the technology to blend with their individual approach to instruction. However, most educators are not fully aware of the benefits that may be obtained by proactively harnessing the available technologies and how they might be able to influence further developments through systematic feedback and suggestions.

- Educational system developers and artificial intelligence (AI) researchers are sometimes unaware of the needs and requirements of typical teachers, with a possible exception of those in the computer science domain. In transferring the notion of a 'user' from the human-computer interaction studies and assigning it to the 'student', the educator's role as the 'implementer/manager/user' of the technology has been forgotten.

The aim of the journal is to help them better understand each other's role in the overall process of education and how they may support each other. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.

The scope of the journal is broad. Following list of topics is considered to be within the scope of the journal:

- Architectures for Educational Technology Systems
- Computer-Mediated Communication
- Cooperative/ Collaborative Learning and Environments
- Cultural Issues in Educational System development
- Didactic/ Pedagogical Issues and Teaching/Learning Strategies
- Distance Education/Learning
- Distance Learning Systems
- Distributed Learning Environments
- Educational Multimedia
- Evaluation
- Human-Computer Interface (HCI) Issues
- Hypermedia Systems/ Applications
- Intelligent Learning/ Tutoring Environments
- Interactive Learning Environments
- Learning by Doing
- Methodologies for Development of Educational Technology Systems
- Multimedia Systems/ Applications
- Network-Based Learning Environments
- Online Education
- Simulations for Learning
- Web Based Instruction/ Training

Editors

Kinshuk, Athabasca University, Canada; Demetrios G Sampson, University of Piraeus & ITI-CERTH, Greece; Ashok Patel, CAL Research & Software Engineering Centre, UK; Reinhard Oppermann, Fraunhofer Institut Angewandte Informationstechnik, Germany.

Editorial Assistant

Barbara Adamski, Athabasca University, Canada.

Associate editors

Nian-Shing Chen, National Sun Yat-sen University, Taiwan; Vladimir A Fomichov, K. E. Tsiolkovsky Russian State Tech Univ, Russia; Olga S Fomichova, Studio "Culture, Ecology, and Foreign Languages", Russia; Piet Kommers, University of Twente, The Netherlands; Chul-Hwan Lee, Inchon National University of Education, Korea; Brent Muirhead, University of Phoenix Online, USA; Erkki Sutinen, University of Joensuu, Finland; Vladimir Uskov, Bradley University, USA.

Advisory board

Ignacio Aedo, Universidad Carlos III de Madrid, Spain; Mohamed Ally, Athabasca University, Canada; Luis Anido-Rifon, University of Vigo, Spain; Gautam Biswas, Vanderbilt University, USA; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bullen, University of British Columbia, Canada; Tak-Wai Chan, National Central University, Taiwan; Yam San Chee, Nanyang Technological University, Singapore; Sherry Chen, Brunel University, United Kingdom; Darina Dicheva, Winston-Salem State University, USA; Michael Eisenberg, University of Colorado, Boulder, USA; Robert Farrell, IBM Research, USA; Brian Garner, Deakin University, Australia; Mark D. Gross, Carnegie Mellon University, USA; Roger Hartley, Leeds University, UK; J R Isaac, National Institute of Information Technology, India; Mohamed Jenni, University of Tunis, Tunisia; Paul Kirschner, Open University of the Netherlands, The Netherlands; William Klemm, Texas A&M University, USA; Rob Koper, Open University of the Netherlands, The Netherlands; Raddy Lelouche, Universite Laval, Canada; David McConnell, Lancaster University, UK; Rory McGreal, Athabasca University, Canada; David Merrill, Brigham Young University - Hawaii, USA; Marcelo Milrad, Växjö University, Sweden; Richiro Mizoguchi, Osaka University, Japan; Permanand Mohan, The University of the West Indies, Trinidad and Tobago; Kiyoshi Nakabayashi, National Institute of Multimedia Education, Japan; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, The University of Electro-Communications, Japan; Thomas C. Reeves, The University of Georgia, USA; Norbert M. Seel, Albert-Ludwigs-University of Freiburg, Germany; David Williamson Shaffer, University of Wisconsin-Madison, USA; Timothy K. Shih, Tamkang University, Taiwan; Yoshiaki Shindo, Nippon Institute of Technology, Japan; Kevin Singley, IBM Research, USA; J. Michael Spector, Florida State University, USA; Timothy Teo, Nanyang Technological University, Singapore; Chiu-Chung Tsaí, National Taiwan University of Science and Technology, Taiwan; Stephen J.H. Yang, National Central University, Taiwan.

Assistant Editors

Sheng-Wen Hsieh, Far East University, Taiwan; Dorota Mularczyk, Independent Researcher & Web Designer; Ali Fawaz Shareef, Maldives College of Higher Education, Maldives; Jarkko Suhonen, University of Joensuu, Finland.

Executive peer-reviewers

http://www.ifets.info/

Subscription Prices and Ordering Information

For subscription information, please contact the editors at kinshuk@ieee.org.

Advertisements

Educational Technology & Society accepts advertisement of products and services of direct interest and usefulness to the readers of the journal, those involved in education and educational technology. Contact the editors at kinshuk@ieee.org.
Abstracting and Indexing


Guidelines for authors

Submissions are invited in the following categories:

- Peer reviewed publications: Full length articles (4000 - 7000 words)
- Book reviews
- Software reviews
- Website reviews

All peer review publications will be refereed in a double-blind review process by at least two international reviewers with expertise in the relevant subject area. Book, Software and Website Reviews will not be reviewed, but the editors reserve the right to refuse or edit review.

For detailed information on how to format your submissions, please see:

http://www.ifets.info/guide.php

Submission procedure

Authors, submitting articles for a particular special issue, should send their submissions directly to the appropriate Guest Editor. Guest Editors will advise the authors regarding submission procedure for the final version.

All submissions should be in electronic form. The editors will acknowledge the receipt of submission as soon as possible.

The preferred formats for submission are Word document and RTF, but editors will try their best for other formats too. For figures, GIF and JPEG (JPG) are the preferred formats. Authors must supply separate figures in one of these formats besides embedding in text.

Please provide following details with each submission: • Author(s) full name(s) including title(s), • Name of corresponding author, • Job title(s), • Organisation(s), • Full contact details of ALL authors including email address, postal address, telephone and fax numbers.

The submissions should be uploaded at http://www.ifets.info/ets_journal/upload.php. In case of difficulties, they can also be sent via email to (Subject: Submission for Educational Technology & Society journal): kinshuk@ieee.org. In the email, please state clearly that the manuscript is original material that has not been published, and is not being considered for publication elsewhere.
Journal of Educational Technology & Society
Volume 12   Number 4   2009

Table of contents

Special issue articles

Guest Editorial – New Directions in Advanced Learning Technologies
Paloma Díaz, Ignacio Aedo Cuevas and Kinshuk

Personalized educational systems

Supporting Teachers in Identifying Students’ Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach
Sabine Graf, Kinshuk and Tzu-Chien Liu

Student Learning Styles Adaptation Method Based on Teaching Strategies and Electronic Media
Ana Lidia Franzoni and Saïd Assar

Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval
Mohamed Koutheair Khribi, Mohamed Jemni and Olfa Nasraoui

Computer supported Collaborative and Social Learning

A Framework for Semantic Group Formation in Education
Asma Ounnas, Hugh C Davis and David E Millard

InstanceCollage: A Tool for the Particularization of Collaborative IMS-LD Scripts
Eloy D. Villasclaras-Fernández, Julio A. Hernández-Gonzalo, Davinia Hernández-Leo, Juan I. Asensio-Pérez, Yannis Dimitriadis, Alejandra Martínez-Monés

Toward a Semantic Forum for Active Collaborative Learning
Yanyan Li, Mingkai Dong and Ronghua Huang

wConnect: Cultivating a Landscape of Online Places for a Developmental Learning Community
Mary Beth Rosson, Hansa Sinha, Dejin Zhao, John Carroll, Craig Ganoe and Jan Mahar

Emerging Web Technologies in Higher Education: A Case of Incorporating Blogs, Podcasts and Social Bookmarks in a Web Programming Course based on Students’ Learning Styles and Technology Preferences
Nauman Saeed, Yun Yang and Suku Sinnappan

New generation of learning technologies

Adaptive Role Playing Games: An Immersive Approach for Problem Based Learning
Pilar Sancho, Pablo Moreno-Ger, Rubén Fuentes-Fernández and Baltasar Fernández-Manjón

Collaborative Paper-based Annotation of Lecture Slides
Jürgen Steimle, Oliver Brdiczka and Max Mühlhäuser

A Methodological Approach to Encourage the Service-Oriented Learning Systems Development
David Diez, Alessio Malizia, Ignacio Aedo, Paloma Díaz, Camino Fernández and Juan-Manuel Dodero

Competence-related Metadata for Educational Resources that Support Lifelong Competence Development Programmes
Demetrios G. Sampson

Project overview

Free Resources for Online Educators - Allan Hancock College

Full length articles

Outdoor Natural Science Learning with an RFID-Supported Immersive Ubiquitous Learning Environment
Tsung-Yu Lui, Tan-Hsu Tan and Yu-Ling Chu
An Investigation into E-Tool Use for Formative Assignment Assessment - Status and Recommendations
Eva Heinrich, John Milne and Maurice Moore
176-192

Multi-Objective Parallel Test-Sheet Composition Using Enhanced Particle Swarm Optimization
Tsu-Feng Ho, Peng-Yeng Yin, Gwo-Jen Hwang, Shyong Jian Shyu and Ya-Nan Yean
193-206

Factors Influencing Adult Learners' Decision to Drop Out or Persist in Online Learning
Ji-Hye Park and Hee Jun Choi
207-217

A Five Dimensional Model for Educating the Net Generation
Ronald Noel Beyers
218-227

Enhancing Eight Grade Students' Scientific Conceptual Change and Scientific Reasoning through a Web-based Learning Program
Ya-Wen Liao and Hsiao-Ching She
228-240

Disabilities and e-Learning Problems and Solutions: An Exploratory Study
Catherine S. Fichten, Vittoria Ferraro, Jennison V. Asuncion, Caroline Chwojka, Maria Barile, Mai N. Nguyen, Ryan Klomp and Joan Wolforth
241-256

Can Online Peer Assessment be Trusted?
L'hadi Bouzidi and Alain Jaillet
257-268

Critical Thinking and ICT Integration in a Western Australian Secondary School
Graham McMahon
269-281

Learner and Instructional Factors Influencing Learning Outcomes within a Blended Learning Environment
Doo Hun Lim and Michael Lane Morris
282-293

Technology's Role in Field Experiences for Preservice Teacher Training
Emily Hixon and Hyo-Jeong So
294-304

The STIN in the Tale: A Socio-technical Interaction Perspective on Networked Learning
Steve Walker and Linda Creanor
305-316

Developing Cognition with Collaborative Robotic Activities
Rubén Mitnik, Miguel Nussbaum and Matías Recabarren
317-330

Are the Functions of Teachers in e-Learning and Face-to-Face Learning Environments Really Different?
Laura Alonso Diaz and Florentino Blázquez Entonado
331-343

The Effects of Mobile Natural-science Learning Based on the 5E Learning Cycle: A Case Study
Tzu-Chien Liu, Hsinyi Peng, Wen-Hsun Wu and Ming-Sheng Lin
344-358

Integration and Deployment of Educational Games in e-Learning Environments: The Learning Object Model Meets Educational Gaming
Javier Torrente, Pablo Moreno-Ger, Iván Martínez-Ortiz and Baltasar Fernandez-Manjon
359-371

Influence of Structure and Interaction on Student Achievement and Satisfaction in Web-Based Distance Learning
Hye-Jung Lee and Iju Rha
372-382

Preservice Teachers' Perceptions of Applets for Solving Mathematical Problems: Need, Difficulties and Functions
Wajeeth Daher
383-395

Book review(s)

Street Gang: The Complete History of Sesame Street (Michael Davis)
Reviewer: Dale Erlandson
396-397
Guest Editorial – New Directions in Advanced Learning Technologies

Paloma Díaz¹, Ignacio Aedo¹ Cuevas and Kinshuk²

¹Computer Science Department, Universidad Carlos III de Madrid, Spain // pdp@inf.uc3m.es // aedo@ia.uc3m.es
²School of Computing and Information Systems, Athabasca University, Canada // kinshuk@ieee.org

ICALT, the International Conference on Advanced Learning Technologies, brings together people who are working on the design, development, use and evaluation of technologies that will be the foundation of the next generation of e-learning systems and technology-enhanced learning environments. In its 8th edition the theme selected was “Learning technologies in the Information society” with a view to encourage the exploration of the role of learning technologies in the transformation from the information society to the knowledge society where everybody (independently of race, sex, abilities, capabilities, ...) can benefit from technologies to enhance her learning process. This opens a world of opportunities for analyzing the use of technology in learning environments that take into account the characteristics and expectations of different kinds of users and different kinds of learning experiences, whether formal or informal, individual or cooperative, life-long or short term. A total of 98 papers were accepted as full papers in the main ICALT conference, what makes a 23.47% acceptance rate. From these papers the editors of this special issue selected a number of contributions that went through a peer review process. We’d like to thank all the reviewers that contributed with their judgment and comments to select the papers in this issue; their names listed below in recognition of their efforts. After the review process, 12 papers were finally selected to illustrate the advances in three relevant areas: Personalized educational systems, Computer supported collaborative and social learning and New generation of educational technologies.

Personalized educational systems are aimed at adapting the educational experience to meet different kinds needs, goals and preferences. This is an important area of research and development where different approaches can be applied to understand how the differences impact the learning process and which kind of tools have to be used, whether computer based or not, to make it more efficient. In this issue two of the papers selected in this category deal with the learning styles of students. Firstly, Sabine Graf, Kinshuk and Tzu-Chien Liu propose in “Supporting Teachers in Identifying Students’ Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach” a tool to identify the learning styles of online students; a valuable information that can assist teachers in understanding how to shape the learning experience for each student. Secondly, in “Student Learning Styles Adaptation Method Based on Teaching Strategies and Electronic Media“, Ana Lidia Franzoni and Said Assar provide a framework to select the best teaching strategy and electronic materials according to the learning style of the student. The third paper is included in this category, “Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval” by Mohamed Koutheair Khribi, Mohamed Jemni and Olfa Nasraoui, deals with collaborative filtering for e-learning recommendations. The authors use web mining techniques and scalable search engine technology to obtain the recommendations from a massive repository of educational resources.

The second group of papers are related to Computer supported Collaborative and Social Learning. CSCL has been a classical topic in education as in many occasions learning occur when interacting with other peers and the mechanisms to create and motivate the groups, as well as the kind of activities and teaching strategies to apply makes up a broad research area. Two papers in this issue deal with group formation. In “A Framework for Semantic Group Formation in Education”, Asma Ounnas, Hugh C. Davis and David E. Millard describe a framework that given a number of constraints defined by the teacher proposes groups based on the use of semantic data about the students features. In “InstanceCollage: a tool for the particularization of collaborative IMS-LD scripts”, Eloy D. Villasclaras-Fernández, Julio A. Hernández-Gonzalo, Davinia Hernández-Leo, Juan I. Asensio-Pérez, Yannis Dimitriadis and Alejandra Martínez-Monés go deeper into the problem of creating efficient groups in the context of IMS-LD. In “Toward a Semantic Forum for Active Collaborative Learning”, Yanyan Li, Mingkai Dong and Ronghui Huang face a different problem: the utility of learning forums. They introduce the concept of semantic forum, where domain ontology and text mining technologies are combined to create semantic link networking on discussion transcripts. CSCL boundaries were enlarged when social online networks became popular, opening up a new universe computer-mediated interactions that might influence formal and, specially, informal learning in many ways. Two papers are concerned with the use of social web technologies with learning purposes. Mary Beth Rosson, Hansa Sinha, Dejin Zhao, John Carroll, Craig Ganoe and Jan Mahar, in “wConnect: Cultivating a Landscape of Online Places for a
Developmental Learning Community”, describe the tools they are using to seed and build a developmental learning community, including extensions of Facebook and design approaches like participatory design. “Emerging Web Technologies in Higher Education: A case of incorporating blogs, podcasts and social bookmarks in a web programming course based on students’ learning styles and technology preferences” by Nauman Saeed, Yun Yang and Suku Sinnappan, describes a survey about students’ learning styles and technology preferences that sets the basis to understand how the use of different emerging technologies can impact students performance.

In the third category in this special issue, New generation of learning technologies, five papers propose innovative techniques to support different processes. Thus, the first paper, “Adaptive Role Playing Games: An Immersive Approach for Problem Based Learning” by Pilar Sancho, Pablo Moreno-Ger, Rubén Fuentes-Fernández, Baltasar Fernández-Manjón, combine techniques from collaborative learning and personalization creating a multiplayer role-playing game that is integrated in a collaborative problem-based learning approach. CoScribe, by Jürgen Steimle, Oliver Brdiczka and Max Mühlhäuser (“Collaborative Paper-based Annotation of Lecture Slides”) is a system that supports students in making collaborative handwritten annotations on printed lecture slides in their Tablet PCs. We move onto the development process with the paper from David Díez, Alessio Malizia, Ignacio Aedo, Paloma Díaz, Camino Fernández and Juan Manuel Dodero (“A methodological approach to encourage the Service-Oriented Learning Systems development”). In this paper, authors make use of Service-Oriented Learning Architecture and domain engineering to encourage the use of learning services. In “Competence-related Metadata for Educational Resources that Support Lifelong Competence Development Programmes”, Demetrios G. Sampson proposes the use of specific meta-data to support the acquisition of abilities through a IEEE LOM Competence-based Application Profile and shows how this approach has been applied in a real project (e-Access2Learn).

With this selection of papers the editors believe to have achieved their main goal: to provide the reader with the state of the art of research in learning technologies as well as a scent of ICALT 2008. We hope you enjoy this special issue and that you explore more contributions to this research area in next ICALT conferences.

Recognizing the contribution of reviewers:
Felix Buendia, Maiga Chang, Sabine Graf, Esther Guerra, Monica Landoni, Tzu-Chien Liu, Antonio Piccinio, Liliana Santa Cruz, Ali Fawaz Shareef, Jarkko Suhonen, Michael Verhaart, Guangbing Yang and Telmo Zarraonandia
Supporting Teachers in Identifying Students’ Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach

Sabine Graf¹, Kinshuk¹ and Tzu-Chien Liu²

¹School of Computing and Information Systems, Athabasca University, Canada // sabine.graf@ieee.org // kinshuk@athabascau.ca
²National Central University, Graduate Institute of Learning and Instruction, Taiwan // ltc@cc.ncu.edu.tw

ABSTRACT

In learning management systems (LMSs), teachers have more difficulties to notice and know how individual students behave and learn in a course, compared to face-to-face education. Enabling teachers to know their students’ learning styles and making students aware of their own learning styles increases teachers’ and students’ understanding about the students’ learning process, allows teachers to provide better support for their students, and has therefore high potential to enhance teaching and learning. This paper proposes an automatic approach for identifying students’ learning styles in LMSs as well as a tool that supports teachers in applying this approach. The approach is based on inferring students’ learning styles from their behaviour in an online course and was developed for LMSs in general. It has been evaluated by a study with 127 students, comparing the results of the automatic approach with those of a learning style questionnaire. The evaluation yielded good results and demonstrated that the proposed approach is suitable for identifying learning styles. DeLeS, the tool which implements this approach, can be used by teachers to identify students’ learning styles and therefore to support students by considering their individual learning styles.

Keywords

Learning styles, Felder-Silverman learning style model, Student modelling, Learning management systems

Introduction

In traditional learning, teachers can easily get an insight into how their students work and learn. However, in online learning, especially when using systems like learning management systems (LMSs), it is more difficult for teachers to see how individual students behave and learn in the system. LMSs such as Moodle (2009), Sakai (2009), and WebCT (2009) are commonly and successfully used in e-learning. They aim at supporting teachers in creating and managing online courses. However, with respect to providing teachers with information about their students, they mainly show how the overall class is using a course rather than focusing on individual students.

This paper focuses on supporting teachers in identifying their students’ learning styles in LMSs. Learning styles can be seen as “a description of the attitudes and behaviours which determine an individual’s preferred way of learning” (Honey & Mumford, 1992, p. 1). Many learning style models exist in literature, such as the learning style model by Kolb (1984), Honey and Mumford (1982), Pask (1976), and Felder and Silverman (1988). While there are still many open issues with respect to learning styles, the learning style models agree that learners have different ways in which they prefer to learn. Furthermore, many educational theorists and researchers consider learning styles as an important factor in the learning process and agree that incorporating them in education has potential to facilitate learning for students.

Knowing students’ learning styles can help in many ways to enhance learning and teaching. First, teachers can benefit by getting information about how their students are used to learn, which provides them with a deeper understanding and might help when explaining or preparing learning material. Furthermore, making students aware of their learning styles and showing them their individual strengths and weaknesses can help students to understand why learning is sometimes difficult for them and is the basis for developing their weaknesses. In addition, students can be supported by matching the teaching style with their learning style. Providing students with learning material and activities that fit their preferred ways of learning can make learning easier for them. This matching hypothesis is supported by many educational theories, as stated and described by Coffield, Moseley, Hall, and Ecclestone (2004). Examples for studies which demonstrated supportive results of this hypothesis include those by Bajraktarevic, Hall, and Fullick (2003) and Graf and Kinshuk (2007).
For considering learning styles in education, the students’ learning styles need to be known first. Brusilovsky (1996) distinguished between two different ways of student modelling: collaborative and automatic. In the collaborative approach, the learners provide explicit feedback which can be used to build and update a student model, such as filling out a learning style questionnaire. In the automatic approach, the process of building and updating the student model is done automatically based on the behaviour and actions of learners while they are using the system for learning. The automatic approach is direct and free from the problem of inaccurate self-conceptions of students. Moreover, it allows students to focus only on learning rather than additionally providing explicit feedback about their preferences. In contrast to learning style questionnaires, an automatic approach can also be more accurate and less error-prone since it analyses data from a specific time span rather than data which are gathered at one specific point of time.

In this paper, we propose an automatic student modelling approach for identifying learning styles in LMSs as well as a tool that implements this approach. Both, the approach and the tool are developed in a generic way and are therefore applicable for LMSs in general. For the theoretical basis regarding learning styles, a well-known and often used learning style model, especially in technology enhanced learning, has been selected, namely the Felder-Silverman learning styles model (FSLSM) (Felder & Silverman, 1988). As mentioned above, many learning style models exist in literature. FSLSM differs from other models since it combines major learning style models like the model by Kolb (1984) and Pask (1976). By combining these models, FSLSM describes learning styles in more detail through characterising each learner according to four dimensions: active/reflective, sensing/intuitive, visual/verbal, and sequential/global. Furthermore, FSLSM considers learning styles as tendencies rather than obligatory types, as done by many other learning style models.

In the following section, the concept for identifying learning styles is introduced. Subsequently, the evaluation of the proposed approach and its results are presented. Section 4 introduces the tool which implements the student modelling approach. Section 5 provides discussion about the proposed approach and the tool, and Section 6 concludes the paper.

A concept for identifying learning styles

The concept for identifying learning styles is based on patterns of behaviour which reveal certain preferences for a particular learning style. The approach is similar to the one of the Index of Learning Styles (ILS) questionnaire (Felder & Soloman, 1997), which is an instrument for identifying learning styles based on the FSLSM. However, while the questionnaire asks students about how they think they prefer to behave and learn, the proposed approach gathers data about how students really behave and learn by observing them during their learning process. The patterns of behaviour which are assumed to be relevant for each learning style dimension are discussed in the next subsection in more detail. Subsequently, the procedure about how to calculate learning styles from the data of the patterns is described.

<table>
<thead>
<tr>
<th>Active/Reflective</th>
<th>Sensing/Intuitive</th>
<th>Visual/Verbal</th>
<th>Sequential/Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>content_visit (-)</td>
<td>content_visit (-)</td>
<td>content_visit (-)</td>
<td>outline_visit (-)</td>
</tr>
<tr>
<td>content_stay (-)</td>
<td>content_stay (-)</td>
<td>ques_graphics (+)</td>
<td>outline_stay (-)</td>
</tr>
<tr>
<td>outline_stay (-)</td>
<td>example_visit (+)</td>
<td>ques_text (-)</td>
<td>ques_detail (+)</td>
</tr>
<tr>
<td>example_stay (-)</td>
<td>example_stay (+)</td>
<td>forum_visit (-)</td>
<td>ques_overview (-)</td>
</tr>
<tr>
<td>selfass_visit (+)</td>
<td>selfass_visit (+)</td>
<td>forum_stay (-)</td>
<td>ques_interpret (-)</td>
</tr>
<tr>
<td>selfass_stay (-)</td>
<td>selfass_stay (+)</td>
<td>forum_post (-)</td>
<td>ques_develop (-)</td>
</tr>
<tr>
<td>selfass_twice_wrong (+)</td>
<td>exercise_visit (+)</td>
<td></td>
<td>navigation_skip (-)</td>
</tr>
<tr>
<td>exercise_visit (+)</td>
<td>ques_detail (+)</td>
<td></td>
<td>navigation_overview_visit (-)</td>
</tr>
<tr>
<td>exercise_stay (+)</td>
<td>ques_facts (+)</td>
<td></td>
<td>navigation_overview_stay (-)</td>
</tr>
<tr>
<td>quiz_stay_results (-)</td>
<td>ques_concepts (-)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>forum_visit (-)</td>
<td>ques_develop (-)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>forum_post (+)</td>
<td>quiz_revisions (+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>quiz_stay_results (+)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Relevant patterns of behaviour

In order to make our approach applicable for LMSs in general, only commonly used features in LMSs were selected to be the basis for the incorporated patterns. These features include: content objects, outlines, examples, self-assessment tests, exercises, and discussion forums. Furthermore, the navigation behaviour of students in the course was considered.

In the next subsections, the characteristics of each learning style with respect to the FSLSM are described and the relevant patterns for identifying learning styles for each dimension are presented, using the literature regarding FSLSM (Felder & Silverman, 1988) as basis. Table 1 provides a summary of the patterns per dimension, using “+” and “-“ for indicating a high and low occurrence of the respective pattern from the viewpoint of an active, sensing, visual, and sequential learning style.

Active/reflective learning style dimension

According to FSLSM, this dimension distinguishes between an active and reflective way of processing information. Active learners learn best by working actively with the learning material, by applying the material, and by trying things out. Furthermore, they tend to be more interested in communicating with others and prefer to learn by working in groups where they can discuss about the learned material. In contrast, reflective learners prefer to think about and reflect on the learning material. Regarding communication, they prefer to work alone.

Identification of active/reflective learning style tendencies

Based on the different preferences regarding communication of active and reflective learners, the forum was used as a relevant feature for identifying students’ learning styles. More specifically, the number of visits in the forum and the number of postings were considered. Based on the characteristics of active and reflective learners described by FSLSM, we can assume that active learners prefer to post more often in order to ask, discuss, and explain issues about the learned material. On the other hand, reflective learners are expected to participate more passively by carefully and frequently reading the postings but only rarely posting by themselves. Other features which help to indicate an active or reflective learning style are self-assessment tests and exercises which allow students to test their knowledge and try things out. A high number of visits of self-assessment tests and exercises as well as spending overall a high amount of time on exercises can give a hint towards an active learning style. On the other hand, due to their preference of reflecting and thinking about the learned material, reflective learners are expected to spend more time on reflecting about the results of their self-assessment tests and exercises as well as spending more time on answering questions of the self-assessment tests. As a consequence, they are expected to answer the same question less often twice wrongly, which is also used as a pattern for identifying a preference for active or reflective learning. Another hint can be gathered from the usage of examples. Active learners prefer to try things out and solve problems by themselves rather than looking at how someone else has solved a problem. Therefore, if students spend only little time on looking at examples, this behaviour can be interpreted as a hint for an active learning style. Furthermore, the usage of content objects and outlines can reveal indications about an active or reflective preference. Reflective learners are expected to visit and spend more time on reading and reflecting about the learning material as well as staying longer at outlines.

Sensing/intuitive learning style dimension

The second dimension of the FSLSM deals with sensing and intuitive learning. Learners with a sensing learning style like to learn facts and concrete learning material, using their sensory experiences of particular instances as a primary source. They like to solve problems with standard approaches and also tend to be more patient with details. Furthermore, sensing learners are considered as more realistic and sensible; they tend to be more practical than intuitive learners and like to relate the learned material to the real world. In contrast, intuitive learners prefer to learn abstract material, such as theories and their underlying meanings, with general principles rather than concrete instances being a preferred source of information. They like to discover possibilities and relationships and tend to be more innovative and creative than sensing learners.
Identification of sensing/intuitive learning style tendencies

For identifying a sensing or intuitive learning style, self-assessment tests and exercises play an important role. Since sensing learners prefer to learn concrete material like facts while intuitive learners prefer to learn abstract material such as theories, the learners’ performance on questions about facts as well as questions about theories and concepts was used as hints for identifying their learning styles. Furthermore, sensing learners prefer to solve problems by applying standard procedures, which they have learned before, and thus like to be well-prepared for tasks such as assignments and exams. Therefore, we can assume that they perform a higher number of self-assessment tests and exercises in order to check their acquired knowledge and practise to apply it. In contrast, intuitive learners tend to be more creative and like to solve problems in new ways. Therefore, a good performance on questions which ask for developing new solutions can give an indication for an intuitive learning style. Furthermore, a basic understanding of the underlying concepts and theories is important for developing new solutions, which can be seen as another argument for intuitive learners to yield on average better results on such questions. Since sensing learners tend to be more patient with details and work carefully but slowly, the time students take for self-assessment tests can be used as another pattern. Since sensing learners also like to check their answers carefully before submitting them, another pattern is the number of revisions performed before submitting a test or exercise. In line with these characteristics, sensing learners are also expected to spend more time on reviewing their results on self-assessment tests and exercises. Furthermore, by looking at students’ performance on questions about details, the sensing learners’ preference for being careful with details can help them in achieving better results on such questions and is therefore used as another pattern. Besides self-assessment tests and exercises, the students’ behaviour regarding content objects and examples can contribute to identifying their preference regarding sensing or intuitive learning. Since sensing learners prefer to learn from concrete material, we can assume that they prefer learning from examples. Therefore, the number of visits and the time spent on examples can act as another pattern. Another argument for the preference for learning from examples for sensing learners is that they like to solve problems based on standard procedures and examples can help them to learn existing approaches for solving problems. On the other hand, intuitive learners are supposed to learn from more abstract material such as content objects and use examples only as supplementary material. Therefore, we can assume that they visit content objects more often and tend to spend more time on them.

Visual/verbal learning style dimension

The visual/verbal dimension deals with the preferred input mode. The dimension differentiates learners who remember best what they have seen (e.g., pictures, diagrams, flow-charts and so on), from learners who get more out of textual representations, regardless of whether they are written or spoken.

Identification of visual/verbal learning style tendencies

A main feature for identifying a visual or verbal learning style is the forum, where verbal learners are expected to use the forum more often and more intensive in order to discuss and communicate with others. Therefore, the number of visits of forum messages, the time spent in the forum, and the numbers of postings are considered as relevant patterns for this dimension. Furthermore, we expect visual learners to perform better on questions about material which was presented in a visual way using, for example, figures and graphics. On the other hand, verbal learners are expected to yield better results on questions about learning material which has been presented in textual format. Another feature which can provide information about students’ preference for either visual or verbal learning is content objects. Based on the preference for textual representation, verbal learners are expected to visit reading material such as content objects more often, which can be used as another pattern.

Sequential/global learning style dimension

In the sequential/global dimension, the learners are characterised according to their understanding. Sequential learners learn in small incremental steps and therefore have a linear learning progress. They tend to follow logical stepwise paths in finding solutions. In contrast, global learners use a holistic thinking process and learn in large leaps. They tend to absorb learning material almost randomly without seeing connections but after they have learned
enough material they suddenly get the whole picture. Then they are able to solve complex problems and put things together in novel ways; however, they have difficulties in explaining how they did it. Because the whole picture is important for global learners, they tend to be more interested in overviews and in a broad knowledge, whereas sequential learners are more interested in details.

Identification of sequential/global learning style tendencies

Regarding identifying students’ preferences for a sequential or global learning style, mainly the performance on specific questions, the usage of outlines and the course overview page, as well as students’ navigational behaviour was considered. Based on sequential learners’ preference for details and global learners’ preference for overviews and seeing the “big picture”, the performance on questions dealing with details and the performance on questions dealing with overview knowledge and connections between concepts were used as patterns. Furthermore, global learners’ interest in relating and connecting topics to each other helps them in interpreting predefined solutions and developing new solutions. Therefore, a better performance on respective questions can be seen as another hint for a global learning style. Another indication for a global learning style with respect to the students’ interest in getting the “big picture” and overview of the learning material can be retrieved from students’ behaviour regarding outlines of chapters and the course overview page. If students use these features often, in terms of visiting and/or spending much time on them, hints for a global learning style can be concluded. Furthermore, students’ navigational behaviour can reveal information about their learning style. According to FSLSM, sequential learners prefer to go through the course step by step, in a linear way, while global learners tend to learn in large leaps, sometimes skipping learning objects and jumping to more complex material. Therefore, the number of skipped learning objects was considered as another pattern.

From behaviour to learning styles

The previous section described the patterns which are incorporated for each dimension as well as whether a high or low occurrence indicates a specific learning style. Based on this information, data about students’ behaviour can be used to get hints for calculating students’ learning styles. For example, if a student visited exercises often, this gives us a hint that the learner prefers an active learning style. Hints \( h_{dim,i} \) are gathered for each dimension \( (dim) \) and each pattern \( (i) \) which includes relevant information for this dimension. Hints are stated by four values: 3 indicates that the student’s behaviour gives a strong indication for the respective learning style (e.g., active), 2 indicates that the student’s behaviour is average and therefore does not provide a specific hint, 1 indicates that the student’s behaviour is in disagreement with the respective learning style and therefore hints for a preference towards the other pole of the learning style dimension (e.g., reflective), and 0 indicates that no information about the student’s behaviour is available. In order to classify the behaviour of students into these four groups, thresholds from literature are used as basis, considering additionally the characteristics of the respective course. For example, such characteristics can include that the discussion forum is mainly used for asking questions to teachers rather than discussing about topics with other students. Another example is that a course may provide printed learning material and/or lectures so that students use content objects mainly for looking up something rather than learning from them.

By summing up all hints and dividing them by the number of patterns that include available information \( (p_{dim}) \), a measure for the respective learning style \( (ls_{dim}) \) is calculated, as stated in formula 1. This measure is then normalised on a range from 0 to 1 as stated in formula 2, resulting in a measure \( nls_{dim} \) where 1 represents a strong positive preference and 0 represents a strong negative preference for the respective learning style. If no pattern includes available information, no conclusion can be drawn.

\[
ls_{dim} = \frac{\sum_{i=1}^{p_{dim}} h_{dim,i}}{p_{dim}} \quad (1)
\]

\[
nls_{dim} = \frac{ls_{dim} - 1}{2} \quad (2)
\]
Evaluation

The proposed student modelling approach was evaluated by a course about object oriented modelling in the Information System and Computer Science curriculum, held at a university in Austria. 127 students participated in the study. However, only data from 75 students were used for the study since we included only data from students who spent more than 5 minutes on filling out the ILS questionnaire, submitted more than half of the assignments, and attended the final exam.

The course included all types of learning objects described in the previous section, namely content objects, outlines, examples, self-assessment tests, exercises, and discussion forums. Moodle (2009) was used for hosting the course. Since Moodle already has a comprehensive tracking mechanism, only one extension was necessary in order to gather data for all above introduced patterns, dealing with tracking how often students are revising their answers in a self-assessment test or exercise. Furthermore, few additional extensions were performed in order to make gathering of data easier, mainly dealing with distinguishing between particular types of learning objects and questions which would also be possible to do without extension but by giving teachers guidelines about how to name particular types of learning objects and questions.

In order to classify the occurrence of behaviour with respect to the investigated patterns, thresholds were used for each pattern. In the next subsection, these thresholds are discussed. Subsequently, the evaluation method and the results of the evaluation are described.

Classifying the occurrence of behaviour

The thresholds for distinguishing whether students have a high, moderate, or low occurrence of behaviour regarding each pattern were determined from literature and adjusted based on the characteristics of the course rather than depending on the students’ average behaviour in the class. Thus, the approach is also applicable for small classes, where the average distribution of learning styles might not apply due to its small size.

The thresholds regarding discussion forum were based on recommendations from Rovai and Barnum (2003), but were lowered since the forum was mainly used for asking questions which were then answered by tutors rather than discussing with other students. Therefore, for the number of visits, thresholds of 7 and 14 visits per week were used, for the time students spent on the forum, thresholds of 5 and 10 minutes per week were used, and for the number of postings, thresholds of 2 and 4 postings per course were used.

Based on the assumptions of García, Amandi, Schiaffino, and Campo (2007), the thresholds for visiting exercises were set to 25% and 75% of available exercises. For self-assessment tests and examples, we used a threshold of 50% and 100% since both types of learning objects were designed in a way that each object might be visited more than once. For outlines, thresholds of 75% and 150% were used. Regarding content objects, students had additionally the possibility to download the learning material for print. Therefore, the content objects were mainly used for looking up information when students were conducting, for example, some exercises or were reflecting about a topic. Therefore, the thresholds for visiting content objects were set to 10% and 20% of all available content objects. Furthermore, the thresholds for visiting the course overview page was determined with 10% and 20% of all visited learning objects.

The thresholds for the time spent on examples, exercises, self-assessment tests, content objects, outlines, and the course overview page were determined as 50% and 75% in relation to the expected learning time of students with high interest in the respective type of learning object, following the recommendation of Garcia et al. (2007). For the time spent on the results of an exercise or self-assessment test, thresholds of 30 seconds and 60 seconds were assumed. Thresholds for the performance of specific question types were assumed as 50% and 75% of correctly answered questions, based on the applied grading system. With respect to revisions of self-assessment tests and exercises, thresholds were determined as 2.5% and 5% of performed self-assessment tests or exercises. The thresholds regarding how often students answered a self-assessment question twice wrong were assumed as 25% and 50% of times a student is asked the same question twice.
Regarding skipping learning objects, we looked at how often students skipped learning objects in relation to the total number of visited learning objects. Thresholds of 1% and 2% of times students used the navigation menu to skip learning objects were assumed.

Method of evaluation

In order to evaluate our approach, its results were compared with the results of the ILS questionnaire (Felder & Soloman, 1997). The proposed approach aims at detecting learning styles for each dimension of the FSLSM on a 3-item scale, distinguishing, for example, between an active, balanced, and reflective learning style. Therefore, the measure introduced above \( nls_{dim} \) was divided into three groups using values of 0.25 and 0.75 as thresholds. These thresholds are based on experiments, showing that using the first and last quarter for indicating learning style preferences for one or the other extreme of the respective dimension and using the second and third quarter for indicating a balanced learning style yields better results than dividing the range into 3 equal parts. Similarly, results of the ILS questionnaire were divided into three groups. For measuring the precision of the proposed approach, including also how close the predicted learning style \( (LS_{predicted}) \) is to the learning style detected by the ILS questionnaire \( (LS_{ILS}) \), the following measure proposed by García et al. (2007) was used:

\[
\text{Precision} = \frac{\sum_{i=1}^{n} \text{Sim}(LS_{predicted}, LS_{ILS})}{n} \cdot 100, \tag{3}
\]

where \( n \) is the number of students. The function \( \text{Sim} \) compares its two parameters \( LS_{predicted} \) and \( LS_{ILS} \) and returns 1 if both are equal, 0.5 if one represents a balanced learning style and the other represents a preference for one of the two poles of the dimension, and 0 if they are opposite.

Results

Table 2 shows the results of the comparison between the proposed approach and the ILS questionnaire. The achieved results range from 73.33% to 79.33%, demonstrating a high precision of the proposed approach for all four dimensions of the FSLSM, and therefore, show that the proposed approach is suitable for identifying learning styles.

<table>
<thead>
<tr>
<th></th>
<th>act/ref</th>
<th>sen/int</th>
<th>vis/ver</th>
<th>seq/glo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>79.33%</td>
<td>77.33%</td>
<td>76.67%</td>
<td>73.33%</td>
</tr>
</tbody>
</table>

DeLeS – A Tool for Detecting Learning Styles

DeLeS stands for “Detecting Learning Styles” and is developed based on the approach for automatically identifying learning styles, introduced in the previous two sections. The tool extends this approach by a user interface which allows teachers to easily specify required information and therefore enables them to detect students’ learning styles, in an automatic way, by using data from students’ behaviour in a course in a LMS. DeLeS is a standalone tool and is developed in a way that it can be used for any LMS. In the next section, the architecture of the tool is described in more detail. Subsequently, the tool is introduced by showing the required and possible user interactions for teachers in order to use the tool.

Architecture

Figure 1 shows the architecture of the tool. The tool consists of two components, the data extraction component and the calculation component.

The data extraction component is responsible for extracting the relevant data from the LMS database in order to calculate learning styles with respect to the four dimensions of FSLSM. The relevant data are based on the above
introduced patterns of behaviour, which are derived from several commonly used features in LMSs. Since one of the main aims of the tool is to be applicable for LMSs in general rather than only for one specific system, heterogeneity of database schemata in different LMSs had to be considered. Therefore, a global schema was built. This can be done by a bottom-up approach, using one LMS database schema as a basis, or by a top-down approach, where the required information acts as a basis. Because the databases of LMSs usually include much more information than required for detecting learning styles, and the database schemata from different LMSs have different structures, the top-down approach was applied. Therefore, a global schema was built where each table in the schema includes data representing one pattern of behaviour. To keep the extraction process as simple as possible, the representation of the data in each table is based on the event-based way data are typically stored in LMS databases. For example, a table includes data about each visit of a learner on examples and stores how much time a student spent at each visit. The overall time a student spent on examples is then calculated automatically by the tool and used as raw data for the calculation component. Figure 2 depicts the process of data extraction.

Figure 1: Tool architecture

Figure 2: Process of data extraction
As a result, the data extraction component delivers raw data which represent the behaviour of the learners regarding the determined patterns. These raw data are then passed to the calculation component, which is responsible for calculating learning styles. The calculation component conducts two steps: First, ordered data are calculated from the raw data based on thresholds which classify the occurrence of behaviour. As argued, for example, by Alberer et al. (2003) and Roblyer and Wiencke (2003), these thresholds can vary from course to course depending, for example, on the structure of the course and the subject. Therefore, the proposed tool recommends thresholds which can be changed by the teachers if necessary. Second, based on the ordered data the learning styles of each student are calculated as described in the proposed approach above and the results are stored as text file.

User interaction for specifying information

In order to use the tool, teachers need to provide some information as well as be able to configure the tool so that it is better adjusted to the characteristics of the investigated course. When starting DeLeS, a configuration file is created, stored in xml format and based on a standard set of information, which includes predefined thresholds, parameters, and settings. The teachers are then asked to provide a name for the configuration file, the name of the LMS as well as information for establishing a connection to the LMS database. For opening the same configuration file at a later point of time, the configuration name and the LMS name need to be entered again.

The created xml file is divided in three parts, including the login data to the LMS database, data about parameters, and data about features and patterns. Data about parameters as well as features and patterns can be specified in the main page of DeLeS (illustrated in Figure 3). Besides the data for the database connection, teachers have to provide some information about parameters, including the number of self-assessment questions, exercises, examples, content objects, and outlines in the respective course. These values are required in order to use meaningful thresholds for the number of visits of these types of learning objects, for example, using 25% and 75% of the performed exercises over all available exercises as thresholds. Furthermore, other parameters, mainly for improving the accuracy of patterns about time spans, can be changed by teachers if they feel the predefined values do not fit their course.

Figure 3: Main page of DeLeS
As can be seen in Figure 3, the main page shows all features and patterns, including a brief description of each pattern. Each pattern can be disabled if the LMS or the course does not provide information regarding the respective pattern. This can be done by clicking on the first symbol behind the respective pattern. After doing so, the pattern is written in grey font and the symbol changes, allowing enabling the pattern again.

For each pattern, thresholds for classifying the occurrence of behaviour are predefined (Graf, 2007). These thresholds can be adjusted in order to fit better the characteristics of the respective course. This can be done by clicking on the last symbol in each line, which leads to a page that shows the current thresholds and allows changing and saving the new ones in the configuration file.

Furthermore, the tool needs to know where the required data for each pattern is stored in the database in order to extract it. One way to provide this information is to use a predefined template where the information about where to find the relevant data is already specified. Such a template is currently available for Moodle version 1.6 and Moodle version 1.4. Another way is to define a new template via the tool. However, for specifying the location of all required data in the database, the teacher (or administrator) needs to be familiar with the database of the LMS. Specifying the location can be done by using the SQL editor, which is reachable via the second symbol in each line. In the first step, the SQL editor shows the required fields for providing information about the respective pattern and asks the teacher to specify which tables are needed to extract the required data. In the second step, three kinds of information are asked: all required fields in the tables for extracting data regarding the pattern, how to connect the tables with each other if more than one table is necessary, including the kind of join as well as the required fields for connecting the tables, and conditions which can optionally be provided and are based on the SQL syntax. After submitting the required information, the editor builds a SQL statement and shows the statement as well as its result when applying it to the database. After confirming, the SQL statement is stored in the configuration file. Besides using the SQL editor, teachers can also write the SQL statement directly in the respective text field on the first page of the editor.

Once all information is specified, the teacher can start the detection process, where the tool extracts data from the LMS database, uses it to calculate learning styles, and stores the results in a text file.

**Discussion**

DeLeS allows detecting learning styles by extracting relevant data about students’ behaviour in an online course from the database of an LMS. The calculation process applies a simple rule for calculating learning styles based on these data, assuming that each pattern of behaviour can give relevant indications for identifying students’ learning styles. The underlying concept of this approach is similar to the concept of a questionnaire, except that in the automatic approach information from students’ behaviour is used rather than asking students about their preferences.

While the calculation procedure of the proposed approach is based on the learning style literature (Felder & Silverman, 1988), related works deal with using a data-driven approach, building models for calculating learning styles from the data about students’ behaviour and data about students’ learning styles gathered by the ILS questionnaire. For example, Cha et al. (2006) used Decision Trees (Dunham, 2002) and Hidden Markov Models (Rabiner, 1989) in order to detect learning styles according to the FSLSM. The results were promising, however, only data from the ILS questionnaire indicating a strong or moderate preference on a specific learning style dimension were considered and data indicating a balanced learning style were excluded. Therefore, further investigations towards a more accurate approach are necessary. Another study was conducted by García et al. (2007), using Bayesian networks (Jensen, 1996) in order to detect learning styles for three dimensions of the FSLSM. They achieved results of 58% for the active/reflective dimension, 77% for the sensing/intuitive dimension, and 63% for the sequential/global dimension, using the same measure as used in this study. Comparing these results to those of our evaluation, it can be seen that the approach presented in this paper achieved similar results for the sensing/intuitive dimension and higher results for the active/reflective and sequential/global dimension.

Apart from using a different concept for calculating learning styles from the behaviour data of students, our approach is different from related work in another important way. While related works aim at identifying learning styles in a specific learning system, our approach is developed for LMSs in general. However, making the approach applicable for LMSs in general is more challenging than developing it only for one specific system. It requires consideration that different LMSs support different features and even if an LMS supports a specific feature and is able to provide
data regarding the patterns related to this feature, teachers also need to use the feature in their courses. Therefore, features in our study were selected in consideration of whether they are commonly available in LMSs and commonly used in online courses. In addition, the procedure for calculating learning styles is developed in a way that it allows missing data in case that information for a pattern cannot be retrieved from the course. However, the more data are available the more accurate the approach is. In this context, our approach includes a relatively high number of patterns of behaviour for each learning style dimension, compared to those of related works, such as the model introduced by García et al. (2007) as well as one of our previous research work (Graf & Kinshuk, 2006). On one hand, a high number of patterns gives more detailed information for identifying learning styles and, on the other hand, it contributes to the general applicability of the approach, considering that some patterns might not be available in all LMSs. Another difference to related works is the tool which has been developed for supporting teachers to use the proposed approach for identifying learning styles in LMSs.

Conclusions

This paper introduced an automatic student modelling approach for identifying learning styles based on the Felder-Silverman learning style model in LMSs as well as a tool that implements this approach and makes it applicable for teachers. The proposed approach uses the behaviour of students while they are learning in order to gather hints about their learning styles. By applying a simple rule-based mechanism, learning styles are calculated based on the gathered indications. The evaluation of the approach yielded good results, showing that the approach is suitable for identifying learning styles with respect to the FSLSM and therefore demonstrating the functionality of the tool.

The proposed approach as well as the tool was developed for LMSs in general rather than for one specific system. LMSs are widely used in online and blended education and therefore, the developed tool is widely applicable and can support many teachers in identifying students’ learning styles. The information about students’ learning styles can be used for providing teachers with more information about their students, showing them that their students have different preferences and ways in which they learn. Furthermore, the information about students’ learning styles can help teachers in understanding why and when students may have difficulties in learning. In addition, the information can be used for making students themselves aware of their own learning styles, helping them to better understand their strengths and weaknesses in the learning process. Furthermore, identifying learning styles in an automatic way is an important step for enhancing adaptivity in LMSs, enabling these systems to present students with courses that fit their learning styles based on their prior behaviour in the course.

Future work will deal with developing a concept for dynamic automatic student modelling, where data from students’ behaviour will be used on the fly for modifying and updating the student model and therefore, allowing the system to immediately respond on students’ needs. Furthermore, future work is planned on improving and evaluating the usability of DeLeS in order to provide teachers with better support.

Acknowledgements

The authors acknowledge the support of National Science Council of the Republic of China, Taiwan, under Contract No. NSC 96-2520-S-008-007-MY2, NSC 097-2811-S-008-001-, and NSC 97-2631-S-008-003-, NSERC, iCORE, Xerox, and the research related gift funding by Mr. A. Markin.

References


Student Learning Styles Adaptation Method Based on Teaching Strategies and Electronic Media

Ana Lidia Franzoni
Instituto Tecnológico Autónomo de México, Computer Science Department, Río Hondo #1, Progreso Tizapan, México, D.F. 01080 // Tel: +52 55 56284063 // Fax: +52 55 56284065 // analidia@itam.mx

Saïd Assar
Télécom & Management Sud Paris, Laboratorie CEMANTIC, 9, rue Charles Fourier 91011, Evry, France // Tel: +33 160764488 // Fax: +33 160764493 // said.assar@it-sudparis.eu

ABSTRACT
Recent research on the learning process has shown that students tend to learn in different ways and that they prefer to use different teaching resources as well. Many researchers agree on the fact that learning materials shouldn’t just reflect the teacher’s style, but should be designed for all kinds of students and all kind of learning styles. Even though they agree on the importance of applying these learning styles to different learning systems, various problems still need to be solved, such as matching teaching content with the student's learning style. In this paper, we describe the design of a personalized teaching method that is based on an adaptive taxonomy using Felder and Silverman's learning styles and which is combined with the selection of the appropriate teaching strategy and the appropriate electronic media. Students are able to learn and to efficiently improve their learning process with such method.

Keywords
Adaptability, electronic media, learning styles, teaching strategies

Introduction
Humans have different ways of learning. Some can assimilate in a better way the knowledge received visually, auditory or through a certain sense. Psychology and cognitive sciences have long-time explored this question. The Dual Coding Theory for example states that information is processed through one of two usually independent channels (Beacham et al., 2002). While one channel processes verbal information such as text or audio, the other one processes visual information like diagrams, images, animations, etc. The Sperry's Nobel Prize winning left-brain / right-brain model of thinking suggested that the right hand side and the left hand side of our brain possessed specialized and differentiated functions (Dervan, et al. 2006). The left cerebral hemisphere is thought to be more verbal, logical or clinical, that is, more analytical, while the right cerebral hemisphere influences more the artistic and the sensing side of our intellectual. Powerful encoding and visualization techniques have shown to enable the creation lasting memory and improve recall. Dual encoding, for example, has proven to be an extremely effective learning tool. The simplest and most common form of which involves presenting the information both textually and visually. “Whole brain” learning is known to be a far more effective way to learn. The better connected the two halves of the brain, the greater the potential of the brain for learning and creativity (Rose, 1998; Dervan, et al. 2006). However, most educational systems have ignored individual differences that exist between learners, such as the learning ability, the background knowledge, the learning goals and the learning style (Ford & Chen, 2001). Educational systems generally provide a unique and standardized teaching material to all learners which tend to benefit to those whose learning style and background knowledge fits well with the teaching material. If the teaching style employed closely matches the student preferred style of acquiring knowledge, learning becomes easier and more natural, results improve and learning time is reduced (Rose, 1998). On the other hand, if for example a student is more visual than verbal and everything is written on the blackboard without auditory resources, student will experience difficulties in attaining the pedagogical goals in the requested time. In few words, traditional teaching material and strategies generally tend to benefit some students more than others.

In this sense, it is necessary to deploy resources to support the learning process in a way that it not only suits the characteristics of a few, but that it adapts to the characteristics of each student. In the context of Information Technology evolution and the availability of large number of electronic media, the idea of matching e-media with appropriate teaching and learning styles has been explored since the late 90’s. There are many studies on the effectiveness of combining multimedia and hypermedia with learning styles in educational systems (Najjar, 1996)...
They attempt to associate specific e-media characteristics to different categories of learners and propose instruments and methods for assessing learning style (Riding & Rayner, 1998). Most of these studies rely on Kolb's Learning Styles Inventory (LSI) (Kolb 1984) and Soloman-Felder Index of Learning Styles (ILS) (Soloman, & Felder, 1993).

However, very few researchers give an idea of which appropriate combinations of electronic media and learning styles are more effective than others. An electronic media can be used in different ways to implement different teaching strategies which can be matched with different learning styles. For example, a discussion forum can be used in different ways. It can be used to assign a practical task to students in such a way that students solve the assigned problem in a collective manner. This fits well with sensitive learning style. The discussion forum can also be used to give a sequential series of theoretical presentations to students who can interact with the teacher. The sequence of presentations associated with the corresponding discussion is an adequate teaching material for sequential style students.

The objective of our research work is the creation of teaching methods and environments that use the vast resources offered by IT in such a way to adapt teaching material and strategies to the learner's skills and learning style. We use Felder & Silverman (1988) model for defining learning style, together with empirically built adaptation taxonomy for matching e-media with combinations of teaching strategies and learning styles. In a previous work, we explored some basic ideas concerning the matching of e-media and learning styles in the context of an experimental e-learning system (Franzoni & Assar, 2007). In this paper, we present a general framework for combining and adapting teaching strategies, learning styles and electronic media. This taxonomy has been experimented in an undergraduate computer science course. First results show that a majority of students have a better assimilation of knowledge and that students appreciated positively the personalized pedagogical material proposed in the course.

The rest of the paper is organized as follows. First, we present related works concerning existing adaptation techniques for adapting learning styles with e-media. Second, we present the pedagogical bases of our work. The third section introduces the adaptive teaching taxonomy with the underlying model. Fourth, the method for applying the adaptive taxonomy is described and illustrated in a teaching experiment. Finally, concluding remarks and comments are presented together with directions for future work.

**Related works**

Recent investigations [(Kwok & Jones 1985), (Carver et al., 1999), (Gilbert & Han 1999), (Grigoriadou, Papanikolaou & Kornilakis 2001), (Stash & De Bra 2004), (Hong & Kinshuk 2004)], try to integrate the learning styles and e-media in the design of their applications. This is not an easy process, however. One of the main difficulties on the designing of hypermedia systems, is linking the learning styles with the hypermedia applications. Most of the teaching systems adaptation that integrates learning styles is based on the premise that adapting the teaching strategies with the students’ learning styles will give better results (Dagger, Wade & Conlan 2003), (Paredes & Rodriguez 2002), (Stern & Woolf 2000), (Triantafillou, Pomportsis & Georgia 2002). Table 1 shows some of the systems found, their learning styles and the type of adaptation.

<table>
<thead>
<tr>
<th>System</th>
<th>Learning style</th>
<th>The adaptation Model</th>
<th>The adaptation behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARTHUR</td>
<td>visual-interactive, auditory-lecture and text styles</td>
<td>The adaptation is achieved by providing different media representations for each learner. Auditory representation is achieved using sounds and streaming audio. To appeal to visual and kinesthetic learners puzzles, animations, drag and drop examples and riddles are used.</td>
<td>Type and usually the sequencing of material they offer based on a framework proposed by the authors</td>
</tr>
<tr>
<td>(Gilbert &amp; Han 1999)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS388</td>
<td>Felder-Silverman learning styles model global-sequential, visual-verbal,</td>
<td>The adaptation is achieved by providing different media representations for each learner. Uses</td>
<td>Based on research studies, (Felder and Silverman, 1988) about the</td>
</tr>
<tr>
<td>System</td>
<td>Features</td>
<td>Adaptation</td>
<td>References</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lane 1999</td>
<td>sensing-intuitive, inductive-deductive styles (Felder &amp; Silverman, 1988)</td>
<td>different types of media such as graphs, movies, text, slideshows</td>
<td>type of instructional material that learners with different learning style prefer</td>
</tr>
<tr>
<td>MANIC (Stern &amp; Woolf 2000)</td>
<td>applies preferences for graphic versus textual information</td>
<td>The adaptation is achieved by providing different media representations for each learner. Uses graphic and textual information</td>
<td>Type and usually the sequencing of material they offer based on a framework proposed by the authors</td>
</tr>
<tr>
<td>INSPIRE (Grigoriadou, Papanikolaou &amp; Kornilakis 2001)</td>
<td>Honey and Mumford categorization of activists, pragmatists, reflectors and theorists based on Kolb (Honey &amp; Mumford, 1992)</td>
<td>The Adaptation lies in presenting a different sequence of alternative contents of the concepts. Concepts can be represented by ‘example’, ‘activity’, ‘theory’, ‘exercise’</td>
<td>Based on research studies (Honey &amp; Mumford, 1992), about the type of instructional material that learners with different learning style prefer (Papanikolaou et al., 2003)</td>
</tr>
<tr>
<td>Tangow (Paredes &amp; Rodriguez 2002)</td>
<td>sensing-intuitive dimension from the Felder-Silverman learning style model (Felder &amp; Silverman, 1988)</td>
<td>The Adaptation lies in presenting a different sequence of alternative contents of the concepts. Concepts can be represented by ‘example’, ‘exposition’</td>
<td>Type and usually the sequencing of material they offer based on a framework proposed by the authors</td>
</tr>
<tr>
<td>AES-CS (Triantafillou, Pomportsis &amp; Georgia 2002) (Triantafillou et al., 2003)</td>
<td>field-dependent (FD) and field-independent (FI) style (Witkin et al., 1977)</td>
<td>Provides field-dependent learners with navigational support tools, such as concept map, graphic path indicator, advanced organizer, in order to help them organize the structure of the knowledge domain. The system guides them through the learning material via adaptive navigation support. Field-independent learners are provided with a learner control option - for them, the system shows a menu from which they can proceed with the course in any order. Learners can switch between different instructional strategies</td>
<td>Adopts several instructional strategies that accommodate learners’ learning style in relation with: the approaches, the control options, the contextual organizers, the study instructions, the feedback, and the lesson structure.</td>
</tr>
<tr>
<td>PHP Programming Course (Hong &amp; Kinshuk 2004)</td>
<td>Active – Reflective, Sensing – Intuitive, Visual – Verbal, Sequential - Global dimension from the Felder-Silverman learning style model (Felder &amp; Silverman, 1988)</td>
<td>The adaptation is achieved by providing different representations for each learner. Uses different types of resources such as concepts, theory, colors, text, slideshows, audio, etc.</td>
<td>Based on research studies, (Felder and Silverman, 1988) about the type of instructional material that learners with different learning style prefer</td>
</tr>
</tbody>
</table>

This review shows that the different adaptation to learning styles systems are done in terms of content adaptation, navigation routes or the use of multiple navigation instruments. However, the election of learning styles seems to be limited, while it is based on the appropriate technology. Also, most of the systems shown, except CS388 and PHP Programming Course, evaluate and adapt to the chosen learning styles dimensions. One disadvantage of CS388 and the PHP Programming Course is that electronic media is limited to graphics, hypertext, audio and video, and that it doesn’t integrate teaching strategies. In this sense, this work is new and significantly different from the previous efforts done by others in the field.
Pedagogic Model

Learning Styles Model by Felder-Silverman

A learning style is defined as the characteristics, strengths and preferences in the way people receive and process information (Felder & Silverman 1988). It refers to the fact that every person has its own method or set of strategies when learning. According to Sewall, there are several theories about learning styles (Sewall, 1986). He did a detailed study of four learning styles evaluation instruments: Myers-Briggs Type Indicator, Kolb’s Learning Style Inventory, Canfield’s Learning Style Inventory and Gregorc’s Type Indicator.

We have selected the Felder and Silverman model as the basis of our taxonomy of adaptive teaching for the following reasons:

- it has been successfully implemented in previous work when individually adapting the electronic learning material (Carver, Howard & Lane, 1999), (Hong & Kinshuk, 2004), (Paredes & Rodriguez, 2002),
- it has been approved by its author and other specialists (Zywno, 2003) (Felder & Spurlin, 2005),
- it is user friendly and the results are easy to interpret,
- the number of dimensions is controlled and can actually be implemented (Paredes & Rodriguez, 2002).

This model rates the student’s learning style in a scale of four dimensions. Each learning style can be defined by answering these four questions:

1. What kind of information does the student tend to receive: sensitive (external agents like places, sounds, physical sensations), or intuitive (internal agents like possibilities, ideas, through hunches)?
2. Through which sensorial channel do the students tend to receive information more effectively: visual (images, diagrams, graphics), or verbal (spoken words, sounds)?
3. How is the information processed: actively (through physical activities and discussions), or reflexively (through introspection)?
4. How does the student make progress: sequentially (with continuous steps), or globally (through leaps and an integral approach)?

Table 2 shows the learning styles dimensions (LSD) that resulted from the latter questions:

<table>
<thead>
<tr>
<th>Learning Style Dimension</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perception (LSD1)</td>
<td>Sensitive (S)</td>
<td>Rather deal with facts, raw data and experiments, they’re patient with details, but don’t like complications</td>
</tr>
<tr>
<td></td>
<td>Intuitive (I)</td>
<td>Rather deal with principles and theories, are easily bored when presented with details and tend to accept complications</td>
</tr>
<tr>
<td>Entry Channel (LSD2)</td>
<td>Visual (Vi)</td>
<td>Easy for them to remember what they see: images, diagrams, time tables, films, etc.</td>
</tr>
<tr>
<td></td>
<td>Verbal (Ve)</td>
<td>Remember what they’ve heard, read or said.</td>
</tr>
<tr>
<td>Processing (LSD3)</td>
<td>Active (A)</td>
<td>Learn by working in groups and handling stuff</td>
</tr>
<tr>
<td></td>
<td>Reflexive (Re)</td>
<td>Learn better when they can think and reflect about the information presented to them. Work better alone or with one more person at most.</td>
</tr>
<tr>
<td>Understanding (LSD4)</td>
<td>Sequential (Seq)</td>
<td>Follow a linear reasoning process when solving problems and can work with a specific material once they’ve comprehended it partially or superficially</td>
</tr>
<tr>
<td></td>
<td>Global (G)</td>
<td>Take big intuitive leaps with the information, may have a difficulty when explaining how they got to a certain result, need an integral vision</td>
</tr>
</tbody>
</table>

Teaching Strategies

Considering that pedagogy includes teaching and learning strategies, we will provide a definition of both. Learning strategies are the strategies used to remember, learn and use information. In this case, responsibility relies on the
student (comprehension and text writing, problem solving, etc.). Students go through a process where they recognize the new knowledge, review previous concepts, organize and restore that previous knowledge, match it with the new one, assimilate it and interpret everything that was seen on the subject.

Teaching strategies (TS) are the elements given to the students by the teachers to facilitate a deeper understanding of the information. The emphasis relies on the design, programming, elaboration and accomplishment of the learning content. Teaching strategies must be designed in a way that students are encouraged to observe, analyze, express an opinion, create a hypothesis, look for a solution and discover knowledge by themselves. Didactic teaching strategy for example refers to an organized and systematized sequence of activities and resources that teachers use while teaching. The main objective is to facilitate the students’ learning.

Among the different components of a teaching strategy, we can mention the way of developing the learning process, and the means and resources used. In this sense, some of the previous studies worth mentioning are for example those of Dunn (1988), who insists on the importance of teaching the students by using methods that adapt to their conceptual preferences. Or Cabrero (2006), who also points out how the applied teaching strategies will take effect on the teaching quality, not only from an individual point of view, but also on the collaboration of the group as a whole.

One crucial aspect of our research is the integration of electronic media with teaching strategies, because of the informational technology breakthroughs that allow us to use a variety of them. On the other hand, we need to link such teaching strategies with the concept of learning styles, something that hasn’t been exploited to the extent that is intended here. The teaching strategies for this case are in table 3.

### Table 3. Teaching strategies (TS)

<table>
<thead>
<tr>
<th>Teaching Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS1 = Games and simulations</td>
</tr>
<tr>
<td>TS2 = Learning based on problem solving</td>
</tr>
<tr>
<td>TS3 = Role playing</td>
</tr>
<tr>
<td>TS4 = Presentation</td>
</tr>
<tr>
<td>TS5 = Discussion panel</td>
</tr>
<tr>
<td>TS6 = Brainstorming</td>
</tr>
<tr>
<td>TS7 = Case study</td>
</tr>
<tr>
<td>TS8 = Question and answer method</td>
</tr>
<tr>
<td>TS9 = Project design method</td>
</tr>
</tbody>
</table>

**Adaptive Teaching Taxonomy**

One usual definition of taxonomy is the following: “*it is the science of orderly classification in natural history*” (Bloom, 1971). Teaching taxonomy seeks to provide the foundations for a classification of the goals within a teaching system (i.e. classification of the desired behavior of the student). From the point of view of the teaching strategies and their definitions, there’s a need to facilitate the implementation of Felder and Silverman’s theories of learning styles by selecting the proper electronic media and teaching strategies for each style (Hong & Kinshuk, 2004), (Carver et al., 1999), (Felder & Silverman, 1988), (Gilbert & Han, 1999), (Paredes & Rodriguez, 2002).

The proposed taxonomy consists on matching the different learning styles with teaching strategies. It also suggests the suitable electronic media as a channel for its representation, thus personalizing it to every student. This taxonomy has been constructed based on our own diverse experiences with Soloman – Felder learning style theory and usage of e-media (Franzoni et al., 2008). It has been checked through an expert panel using the Delphi method which was held during the *III Congreso de Estilos de Aprendizaje* at Cáceres (Spain) in July 2008.

This taxonomy is based on the four learning styles dimensions (LSD):

\[
\text{LSD} = \{\text{LSD}_1, \text{LSD}_2, \text{LSD}_3, \text{LSD}_4\}
\]
Each dimension is defined as a combination of four values according to the learning styles dimension (see table 2):
\[ \text{LSD}= \{(\text{Sensitive (S)} / \text{Intuitive (I)}), (\text{Visual (Vi) / Verbal (Ve)}), (\text{Active (A) / Reflexive (Re)}), (\text{Sequential (Seq) / Global (G)})\}. \]

In this case, there are 16 \((2^4)\) learning styles combinations (LSC):
\[ \text{LSC}=\{(S,Vi,A,Seq), (S,Vi,A,G), (S,Vi,R,Seq), (S,Vi,R,G), (S,Ve,A,Seq), (S,Ve,A,G), (S,Ve,R,Seq), (S,Ve,R,G), (I,Vi,A,Seq), (I,Vi,A,G), (I,Vi,R,Seq), (I,Vi,R,G), (I,Ve,A,Seq), (I,Ve,A,G), (I,Ve,R,Seq), (I,Ve,R,G)\}. \]

A detailed review of the learning styles theory helped us establish the following three attributes for a learning style:
- description,
- appropriate pedagogical method
- characteristics of the media to be used

Then, the description and the appropriate method were associated to find the suitable teaching strategies, and finally the characteristics of media to be used mentioned in the learning styles theory was linked with the electronic media.

Each LSC can be associated with a teaching strategy (TSi):
\[ \text{TS} = \{\text{TS}_1, \text{TS}_2, \ldots , \text{TS}_9\}. \]

Teaching strategies hold a one-to-many relationship with the learning styles. There can be one or many teaching strategies that accommodate one learning style.

Each LS can be associated with an appropriate electronic media (EMi):
\[ \text{EM} = \{\text{EM}_1, \text{EM}_2, \ldots , \text{EM}_{27}\}. \]

Learning style hold a one-to-many relationship with the electronic media. For each learning style, there are one or many teaching strategies that can be implemented by one or many electronic media based on adequate learning style (see Figure 1).

![Figure 1. Adaptive Teaching Taxonomy relation entity diagram](image)

For example, student with a sensitive style prefer practical content and methods that allow the solution of problems. When selecting the appropriate strategies, these must give priority to such practical work (learning based on problems, presentation and question and answer method). Electronic media, in turn, should contribute to these
priorities, animations, simulation and forums (among others), are examples of media that allow the implementation of pragmatic solutions that can also be based on problems.

Considering the intuitive style, as it prefers concepts, an exposition strategy is used, even though a discussion panel can also meet the requirements. On the other hand, considering that intuitive persons are innovators, a role playing session can also meet the objective. If we associate these with the electronic media, we find that student response systems, Internet research and webquest (among others), allow us to adapt these strategies. Tables 3, 4, 5 and 6 present in detail the adaptive teaching taxonomy for each learning style dimension. Integration elements for LSD₁ [Perception (Sensitive, Intuitive)] are shown on Table 3.

Table 3. Adaptive taxonomy for the LSD₁ [Perception (Sensitive, Intuitive)]

<table>
<thead>
<tr>
<th>Perception Specifications</th>
<th>Sensitive</th>
<th>Intuitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Practical. Don’t like courses without an immediate link to the real world</td>
<td>Conceptual, innovative, oriented to theory and meaning, enjoy working with abstract problems and mathematic formulations.</td>
</tr>
<tr>
<td>Appropriate pedagogical method</td>
<td>Specific, facts and procedure oriented, enjoy problem solving by following well established procedures, patient when dealing with details, enjoy practical work, lab class and can memorize things easily</td>
<td>Are innovative and hate repetitive work, rather discover possibilities and relationships, assimilate new concepts easily, don’t like courses that require much memory and tedious calculation.</td>
</tr>
<tr>
<td>Characteristics of the media to be used</td>
<td>Practical, problem solving oriented, laboratory and experiments</td>
<td>Theoretical, abstraction and math related</td>
</tr>
<tr>
<td>Associated Teaching Strategies</td>
<td>Presentation</td>
<td>Discussion panel</td>
</tr>
<tr>
<td></td>
<td>Question and answer method</td>
<td>Games and simulations</td>
</tr>
<tr>
<td></td>
<td>Learning based on problem solving</td>
<td>Role playing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Project design method</td>
</tr>
</tbody>
</table>

Integration elements for LSD₂ [Entry Channel (Visual, Verbal)] are shown on Table 4.

Table 4. Adaptive taxonomy for the LSD₂ [Entry Channel (Visual, Verbal)]

<table>
<thead>
<tr>
<th>Entry Channel Specifications</th>
<th>Visual</th>
<th>Verbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Highly visual elements</td>
<td>Oral and text elements</td>
</tr>
<tr>
<td>Appropriate pedagogical method</td>
<td>Rather work with visual representations when receiving information and remember what they see</td>
<td>Rather receive information spoken or verbally and remember what they read or hear</td>
</tr>
<tr>
<td>Characteristics of the media to be used</td>
<td>Visual representations and diagrams</td>
<td>Text and sounds</td>
</tr>
<tr>
<td>Associated Teaching Strategies</td>
<td>Games and simulations</td>
<td>Discussion panel</td>
</tr>
<tr>
<td></td>
<td>Presentation</td>
<td>Brainstorming</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Question and answer method</td>
</tr>
</tbody>
</table>

Integration elements for LSD₃ [Processing (Active, Reflexive)] are shown on Table 5.

Table 5. Adaptive taxonomy for the LSD₃ [Processing (Active, Reflexive)]

<table>
<thead>
<tr>
<th>Processing Specifications</th>
<th>Active</th>
<th>Reflexive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Applicable and group work</td>
<td>Write short summaries</td>
</tr>
<tr>
<td>Appropriate</td>
<td>Tend to comprehend and assimilate new</td>
<td>Think about quietly before go ahead</td>
</tr>
</tbody>
</table>
Integration elements for LSD₄ [Understanding (Sequential, Global)] are shown on Table 6.

**Table 6. Adaptive taxonomy for the LSD₄ [Understanding (Sequential, Global)]**

<table>
<thead>
<tr>
<th>Understanding Specifications</th>
<th>Sequential</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Orderly, step by step and sequential</td>
<td>See everything as a whole</td>
</tr>
<tr>
<td><strong>Appropriate pedagogical method</strong></td>
<td>Learn through small orderly steps when these are logically associated and follow small orderly steps logically associated when solving problems</td>
<td>Learn through big leaps, suddenly and almost randomly, can solve complex problems quickly and put things together in an innovative way may have difficulties to explain how they did it</td>
</tr>
<tr>
<td><strong>Characteristics of the media to be used</strong></td>
<td>That allows content to be shown in steps (chapters)</td>
<td>That allow to see everything as a whole</td>
</tr>
<tr>
<td><strong>Associated Teaching Strategies</strong></td>
<td>Presentation Question and answer method</td>
<td>Role playing Brainstorming Case study Project design method</td>
</tr>
</tbody>
</table>

Table 7 show how can helping teachers, through knowledge of the ways our students learn to solve the problem of integrating new information technologies and configure new teaching and learning situations.

Knowing the learning styles of students we will select the type of teaching strategies (see table 7) and the most appropriate electronic media. (See table8)

**Table 7. Adaptive taxonomy: LS dimensions and TS relationships**

<table>
<thead>
<tr>
<th>Teaching strategy</th>
<th>Games and simulations</th>
<th>Learning based on problem solving</th>
<th>Role playing</th>
<th>Presentation</th>
<th>Discussion panel</th>
<th>Brainstorming</th>
<th>Case study</th>
<th>Question and answer method</th>
<th>Project design method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning styles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitive</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intuitive</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Reflexive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Sequential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
According to teaching strategies selected in the table above the teacher can select the specific electronic media that should be the best appropriate to the instruct knowledge based on the learning style (see Table 8).

**Table 8. Adaptive taxonomy: LS dimensions and EM relationships**

<table>
<thead>
<tr>
<th>Electronic media</th>
<th>Learning styles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitive</td>
</tr>
<tr>
<td><strong>Audio</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Audio Recording</td>
<td>X</td>
</tr>
<tr>
<td>Audioconference</td>
<td>X</td>
</tr>
<tr>
<td><strong>Collaboration</strong></td>
<td></td>
</tr>
<tr>
<td>Forums</td>
<td>X</td>
</tr>
<tr>
<td>Online learning communities</td>
<td>X</td>
</tr>
<tr>
<td>Weblog or blog</td>
<td>X</td>
</tr>
<tr>
<td>Wikis</td>
<td>X</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td></td>
</tr>
<tr>
<td>Chat (Messenger)</td>
<td>X</td>
</tr>
<tr>
<td>e-mail</td>
<td>X</td>
</tr>
<tr>
<td><strong>Diagrams</strong></td>
<td></td>
</tr>
<tr>
<td>Animations</td>
<td>X</td>
</tr>
<tr>
<td>Graphics</td>
<td>X</td>
</tr>
<tr>
<td>Pictures</td>
<td>X</td>
</tr>
<tr>
<td>Simulations</td>
<td>X</td>
</tr>
<tr>
<td><strong>Read</strong></td>
<td></td>
</tr>
<tr>
<td>Digital magazines</td>
<td>X</td>
</tr>
<tr>
<td>Digital newspapers</td>
<td>X</td>
</tr>
<tr>
<td>eBooks</td>
<td>X</td>
</tr>
<tr>
<td>Hypertext (web pages)</td>
<td>X</td>
</tr>
<tr>
<td>Slideshows</td>
<td>X</td>
</tr>
<tr>
<td><strong>Search</strong></td>
<td></td>
</tr>
<tr>
<td>Internet research</td>
<td>X</td>
</tr>
<tr>
<td><strong>Tutoring</strong></td>
<td></td>
</tr>
<tr>
<td>Course Legacy System</td>
<td>X</td>
</tr>
<tr>
<td>Student Response System</td>
<td>X</td>
</tr>
<tr>
<td>Tutorial systems</td>
<td>X</td>
</tr>
<tr>
<td>WebQuest</td>
<td>X</td>
</tr>
<tr>
<td><strong>Video</strong></td>
<td></td>
</tr>
<tr>
<td>Podcast</td>
<td>X</td>
</tr>
<tr>
<td>Recorded live events</td>
<td>X</td>
</tr>
<tr>
<td>Videoconference</td>
<td>X</td>
</tr>
<tr>
<td>Videos</td>
<td>X</td>
</tr>
<tr>
<td>Web seminars (broadcasts)</td>
<td></td>
</tr>
</tbody>
</table>

**Guidelines for use**

The analysis of tables 7 and 8 should allow the teacher to determine the most appropriate teaching strategy and course material. Different approaches can be used. A recommendable approach consists in clustering students with similar learning styles and using the appropriate teaching strategy and material for each of the groups. Usually, the teacher is not able to implement such an approach, due for example to course time constraints, unavailability of the appropriate resources, etc. Should this be the case, another plausible approach consists of the identification of the “group average style” and the selection of the material accordingly. A third alternate approach (and perhaps the most recommendable one, should the resources allow it) consists of the use of different types of materials (thus targeting different styles) for a set of two or three learning units at a time. The selected material would be used on a rotational basis. This can be done with the integration of teams or groups of students having different learning styles. The adoption of this third approach allows the creation of team group skills for the students. Since the analysis of the table can result in having a list of suggestions (on teaching strategies to employ) that is still too long/complex to really serve as a guide for the teacher. In this situation the teacher might want to focus only on the teaching strategy that is representative of each category of learning style. This is illustrated in the following, overall recommendations are presented to select teaching strategy and prepare e-media material for each learning style.
Sensitive Learning Style: The content must be practical, courses must have an immediate connection with the real world, using concrete methods that are oriented towards facts and procedures that follow previously established techniques. The requested homework must be detailed, not global, including problem solving, laboratory exercises and concept memorization.

   Teaching Strategy: Problem solving based learning.
   Electronic Media: Forums

Intuitive Learning Style: The content must be innovative, oriented to theory and meanings, with abstractions and mathematical formulae, avoiding repetitive methods. The requested homework must include the discovery of relations and actions. The introduction of new concepts can be used but not as memorizing facts but as abstractions. Teaching Strategy: Discussion Panel

   Electronic Media: Forums, Wikis, E-mail

Visual Learning Style: The content must be a heavy on visual components. The requested homework must include actions to visualize, the information gathering must use visual representations, images must be used in order to make it easier for the students to remember the contents, and the teacher can request diagrams that summarize the homework.

   Teaching Strategy: Simulations and games
   Electronic Media: Electronic Presentations, Videos (pedagogic), Animations

Verbal Learning Style: The content must have a lot of oral and textual components. The requested homework must include written essays or oral presentations, the information gathering must use textual representations, texts must be used in order to make it easier for the students to remember the contents, and the teacher can request abstracts that summarize the homework.

   Teaching Strategy: Brainstorm
   Electronic Media: Chats, blog, Forums

Active Learning Style: Students tend to comprehend and assimilate new information when they practice using it (discussion, implementation, group presentations) and rather learn working with others. The content must be applicable. The requested homework must include work in groups.

   Teaching Strategy: Role playing
   Electronic Media: Electronic Presentations, Digital Magazines, Digital Newspapers

Reflexive Learning Style: Students observe and ponder experiences. Data are collected and analyzed thoroughly about before any conclusion is made. The content must be related with experiences. The requested homework must include personal work.

   Teaching Strategy: Cases study
   Electronic Media: E-books

Sequential Learning Style: The content must be written orderly, step by step. The requested homework must consist of small orderly steps that are logically associated to the problems being solved. This allows content to be shown in steps (chapters).

   Teaching Strategy: Presentation
   Electronic Media: audioconference

Global Learning Style: The content must be written in big leaps, suddenly and almost randomly. Students can solve complex problems quickly and put things together in an innovative way but may have difficulties to explain how they did it. This allows seeing everything as a whole.

   Teaching Strategy: Project design method
   Electronic Media: Internet research

Application Method

The adaptive teaching taxonomy is a suggestion to change our way of teaching, student-centered, knowing the meaning and practical applications of the theory of learning styles. It’s not that each of our actions conform to the
teaching styles of student learning. For applying the method in this example this would be impossible because of the diversity of styles found in the classroom. It is that throughout our classes offer a clear attention to the students of different learning styles, we organize activities that take into account individual learning style, that assessments are designed taking into account the variety of learning styles of students, by this way the student learning can guide properly.

The proposed method will link learning styles, teaching strategies and electronic media based on the adaptive taxonomy described above, according to a study plan or particular educational objectives to support teaching. There are three factors that affect the application method: teacher, student and the method itself.

There are three main phases for this method (see Figure2):
- the application of the Felder and Silverman learning styles survey, when it’s a traditional class, like the example shown in this article, as mentioned before, it is impossible to generate all the material in all styles of learning, therefore it is necessary to take into account the predominant styles of the students who will be the basis for the two examples selected targets of the material to cover,
- the study plan is reviewed to set the course objectives,
- selection of the teaching strategies and electronic media is according to the adaptive teaching taxonomy and based on the results of the learning styles questionnaire.

Not all the resulting teaching strategies and corresponding electronic media have to be used for every style. You can select the ones that are available.

**Example of method application**

The Instituto Tecnológico Autónomo de México’s (ITAM) Algorithms and Programs course (ID course - COM11101) for first year engineering students was used to test the validity of the method.

In the first phase the Felder and Silverman learning styles questionnaire was given to 26 students. The results are shown in Table 9.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Learning Style</th>
<th>Percentage of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSD₁ Perception</td>
<td>Sensitive</td>
<td>62%</td>
</tr>
<tr>
<td></td>
<td>Intuitive</td>
<td>38%</td>
</tr>
<tr>
<td>LSD₂ Entry Channel</td>
<td>Visual</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Verbal</td>
<td>15%</td>
</tr>
<tr>
<td>LSD₃ Processing</td>
<td>Active</td>
<td>62%</td>
</tr>
<tr>
<td></td>
<td>Reflexive</td>
<td>38%</td>
</tr>
<tr>
<td>LSD₄ Understanding</td>
<td>Sequential</td>
<td>62%</td>
</tr>
<tr>
<td></td>
<td>Global</td>
<td>38%</td>
</tr>
</tbody>
</table>

The results show that in the Perception dimension, students are more sensible than intuitive. The most significant difference is in the Entry Channel dimension, where 85% students came out to be visual and only 15% were verbal. In the case Processing dimension, it was found that most of the students are active. Finally, in the Understanding dimension, it shows that they are mostly sequential. As a result, the predominant combination for each dimension style of the class is \{(Sensitive/Visual/Active/Sequential)\}.
In the second phase, the study plan is reviewed to specify the objectives of the course. The following information was used: The main objective of this prominently formative course is to develop within the student the ability to analyze and solve problems in a methodic way, as well as to express its solution in algorithmic terms. The student will get to know the basic techniques of procedural programming and will use them through C programming language.

This course is complemented with a two hour lab practice every other Friday. During this lab class the teacher will guide the students towards the solution of problems using “algorithms and programs”.

Considering this course’s characteristics, it is highly recommended that students solve the largest number of possible problems. Practice and application on specific problems of the concepts seen in class is what will allow a student to fully understand the theory.

These are the topics of the full course “Algorithms and Programs”, which are arranged based in the specific objectives. Here are the different modules: 1. Course Presentation. 2. Algorithms, Programs and Flowcharts. 3. C Programming Language. 4. Modular Programming. 5. Unidimensional Array. 6. Bidimensional Array. 7. File and Character Chains. 8. Introduction to Structures. In this case, topic number 2 “Algorithms, Programs and Flowcharts” are the base for our example.

In this third phase, a selection of the teaching strategies and electronic media based on the adaptive teaching taxonomy is done according to the results of the learning styles questionnaire, and also based on the predominant style at this class {(Sensitive/Visual/Active/Sequential)}. Table 10 explain the description for each learning style selected, the appropriate pedagogical method, the characteristics of the media to be used, the teaching strategies and the fitting electronic media to be used for this particular style. The description give details about the preferences for the sensitive style, like must be practical, the material must be linked to the real world, with a highly visual approach and easily applicable; teamwork must be encouraged too, the teacher needs to consider all of this for create the course material. Tables 11 and 12 describe the example, how to consider this in the syllabus (Algorithms, Programs and Flowcharts topic).

<table>
<thead>
<tr>
<th>Perception Specifications</th>
<th>Sensitive</th>
<th>Visual</th>
<th>Active</th>
<th>Sequential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Practical. Don’t like courses without an immediate link to the real world</td>
<td>Highly visual elements</td>
<td>Applicable and group work</td>
<td>Orderly, step by step and sequential</td>
</tr>
<tr>
<td>Appropriate pedagogical method</td>
<td>Specific, facts and procedure oriented, enjoy problem solving by following well established procedures, patient when dealing with details, enjoy practical work, lab class and can memorize things easily</td>
<td>Rather work with visual representations when receiving information and remember what they see</td>
<td>Tend to comprehend and assimilate new information when they practice using it (discussion, implementation, group presentations) and rather learn working with others</td>
<td>Learn through small orderly steps when these are logically associated and follow small orderly steps logically associated when solving problems</td>
</tr>
<tr>
<td>Characteristics of the media to be used</td>
<td>Practical, problem solving, laboratory and experiments</td>
<td>Visual representations and diagrams</td>
<td>Group work and cooperation</td>
<td>That allows content to be shown in steps (chapters)</td>
</tr>
<tr>
<td>Teaching Strategies</td>
<td>Presentation Question and answer method Learning based on problem solving</td>
<td>Games and simulations Presentation</td>
<td>Games and simulations Learning based on problem solving Role playing Discussion panel Brainstorming Project design</td>
<td>Presentation Question and answer method</td>
</tr>
</tbody>
</table>
According to electronic media categories selected in the table above the teacher can select the specific electronic media that should be the best appropriate to instruct knowledge (see Table 8).

<table>
<thead>
<tr>
<th>Table 11. Using adaptive taxonomy for algorithms topic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific Objective</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
</tbody>
</table>
| 1. Identify places where algorithms will be used. | - Set examples using algorithms. | *Learning based on problem solving  
*Brainstorming | Communication(chats, email), Collaboration (forums, wikis) |
| 2. Define Algorithms. | - Define algorithms using previous examples  
- Establish problems to be solved using algorithms | * Role playing  
*Question and answer method  
*Discussion panel | Collaboration (forums, wikis)  
Internet research |
| 3. Solve simple problems using algorithms. | - Solve a problem using algorithms | * Learning based on problem solving | Diagrams(animations, graphics, pictures) |

<table>
<thead>
<tr>
<th>Table 12. Using adaptive taxonomy for flowcharts topic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific Objective</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>1. Define flowcharts.</td>
</tr>
</tbody>
</table>
| 2. Compare algorithms with flowcharts. | - Identify the pros of solving problems using flowcharts.  
- Establish the importance of using diagrams to help solving more complex problems. | * Presentation  
*Question and answer method | Read(eBooks, hypertext (web pages), slideshows) |

The different ways of doing the course helps in raising the learning abilities of the student. The teacher displays the information in different ways, with different resources, making the learning process easier due to the fact that some people are more receptive to some kind of information than the others. The previous table displays the wide variety of resources that the teacher might use (if available), according to the course objectives. It is also helpful to identify and select the different tools that might be used. Examination of the table allows the teacher to get a better knowledge of the different potentially useful tools, as well. This can result in a better integration of strategies and selection of instructional tools (some of which might be unknown by the teacher) in later courses. The results of the students were better than other courses which did not use our method. The student's evaluation of the course was much better as well. The use of several resources helps all kinds of students; they were globally very satisfied with the electronic media used. The method was accepted positively by the students. They thought that this is an innovative idea that can help people have a better performance whilst learning.

There are many studies concerning learning styles and their usage in teaching methods, and there are many tutoring systems without a pedagogical method (Gilbert et al., 2005). This adaptive teaching taxonomy presented here is different because it deals directly with the problem of matching teaching strategies with electronic media based on learning styles. Even though the presented example is a traditional class (face to face), it can be completely automated if the discussed set of rules is applied. This way, if an e-learning system is implemented, full personalization of the learning process may be achieved.
Conclusions

The work presented in this paper describes the development of an integrated taxonomy combining learning styles, different teaching strategies and the corresponding appropriate electronic media. The objective is to provide a structured method to help in facilitating the learning process and personalizing the pedagogical resources. This method can be used in traditional face to face classes where the teacher can calculate the course’s student's representative learning style to choose the suitable media as proposed by the adaptive teaching taxonomy. It can also be used in distance learning courses where it acts like a catalyst to achieve an automatic personalization in the hypermedia systems. It is worth mentioning that combining teaching strategies with electronic media as proposed by our method doesn’t act in an excluding way. It can be combined with any additional teaching approach and/or teaching resources. Because the method and the adaptive taxonomy are user friendly, the person implementing this method doesn’t have to be information technology subject matter expert.

We consider the recommendations on teaching strategies and electronic media that match a certain learning style as an important contribution to the field of pedagogical teaching methods. The evaluation of student's learning style gives a strong insight about the students' ability to capture the teacher's message. The proposed taxonomy offers a wide range of possibilities for building a course. Even if full personalization is not possible in face to face teaching, the teacher can develop different versions of the teaching material so that to fit to the learning styles of the largest number of students. It might happen that a teacher does not know the students’ learning styles. Furthermore, he/she might not know either the appropriate educational strategies or instructional material for their courses. The presented taxonomy is thus a useful tool to get a better knowledge of the wide variety of resources available to use in class.

A two phase evaluation of the method to test its efficiency is actually under investigation. The first phase will deal with available off the shelf educational software, and the second phase will be in a suited system under implementation at the ITAM. The objective is to test full personalization of the learning process and to validate the adaptive taxonomy in a quantitative based research approach. Another direction of research is to test the adaptive taxonomy in different knowledge domains. We tend to believe that for a given learning style, some strategies and electronic media apply differently when teaching different domains. Specific adaptive taxonomy combinations can be found to be more useful for certain knowledge domains like linguistics, programming or physics for example.

References


Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval

Mohamed Koutheaïr Khribi¹, Mohamed Jemni¹ and Olfa Nasraoui²

¹Technologies of Information and Communication Lab, Higher School of Sciences and Technologies of Tunis, University of Tunis, Tunisia // mk.khribi@uvt.rnu.tn // mohamed.jemni@fst.rnu.tn
²Knowledge Discovery & Web Mining Lab, University of Louisville, USA // olfa.nasraoui@louisville.edu

ABSTRACT
In this paper, we describe an automatic personalization approach aiming to provide online automatic recommendations for active learners without requiring their explicit feedback. Recommended learning resources are computed based on the current learner’s recent navigation history, as well as exploiting similarities and dissimilarities among learners’ preferences and educational content. The proposed framework for building automatic recommendations in e-learning platforms is composed of two modules: an off-line module which pre-processes data to build learner and content models, and an online module which uses these models on-the-fly to recognize the students’ needs and goals, and predict a recommendation list. Recommended learning objects are obtained by using a range of recommendation strategies based mainly on content based filtering and collaborative filtering approaches, each applied separately or in combination.

Keywords
E-learning, Automatic Personalization, Recommender Systems, Content based filtering, Collaborative Filtering

Introduction
Up to the very recent years, most e-learning systems have not been personalized. Several works have addressed the need for personalization in the e-learning domain. However, even today, personalization systems are still mostly confined to research labs, and most of the current e-learning platforms are still delivering the same educational resources in the same way to learners with different profiles. In general, to enable personalization, existing systems used one or more types of knowledge (learners’ knowledge, learning material knowledge, learning process knowledge, etc). Generally, personalization in e-learning systems concerns: adaptive interaction, adaptive course delivery, content discovery and assembly, and adaptive collaboration support. The category of adaptive course delivery represents the most common and widely used collection of adaptation techniques applied in e-learning systems today. Typical examples include dynamic course re-structuring and adaptive selection of learning objects, as well as adaptive navigation support, which have all benefited from the rise of using recommendation strategies to generate new and relevant links and items. In fact, one of the new forms of personalization in e-learning environment is to give recommendations to learners in order to support and help them through the e-learning process.

A number of personalized systems have relied on explicit information given by a learner (demographic, questionnaire, etc) and have applied known methods and techniques of adapting the presentation and navigation (Chorfi et al., 2004). As explained in (Brusilovsky, 1996), two different classes of adaptation can be considered: adaptive presentation and adaptive navigation support. Later, in (Brusilovsky, 2001), the taxonomy of adaptive hypermedia technologies was updated to add some extensions in relation with new technologies. Then, the distinction between two modes of adaptive navigation support became a necessity, especially with the growth of recommender systems. Automatic recommendation implies that the user profiles are created and eventually maintained dynamically by the system without explicit user information. Examples include amazon.com’s personalized recommendations and music recommenders like Mystrand.com in commercial systems (Mobasher 2006), smart recommenders in e-learning (Zaiane, 2002), etc. In general, such systems differ in the input data, in user modeling strategies, and in prediction techniques. Several approaches for automatic personalization have been reported in the literature, such as content-based or item-based filtering, collaborative filtering, rule-based filtering, and techniques relying on Web usage mining, etc (Nasraoui, 2005). Web recommender systems can be categorized depending on these approaches. Content-based filtering (or item-based filtering) systems recommend items to a given user based on the correlation between the content of these items and the preferences of the user (Meteren et al., 2000). This means that the recommended items are considered to be similar to those seen and liked by the same user in the past. Thus, there is no notion of a community of users, rather only one user profile is considered while...
making recommendations. Classical examples of systems applying content based filtering approach include among other Personal webwatcher (Mladenic, 1996), syskill and webert (Pazzani et al., 1997), etc.

Collaborative filtering system recommends items that are liked by other users with similar interests. Thus, the exploration of new items is assured by the fact that other similar user profiles are also considered. Examples of such systems include GroupLens (Konstan et al., 1997) and (Sarwar et al., 1998). Hybrid recommender systems combine several recommendation strategies to provide better performance than either strategy alone. Most hybrids work by combining several input data sources or several recommendation strategies. There are many hybridization methods reported in the state of the art, content/collaborative hybrids are the most popular hybrid strategies. Generally, web recommender systems tend to use web mining techniques in one or more stage of the recommendation process. In e-learning, the interest in using web mining has recently increased, especially with the rapid spread of web based learning environments in education and the growing need to give personalized services to students. In e-learning systems, web mining techniques are used to learn all available information about learners and build models to apply in personalization. A detailed description about using and applying educational data mining was given in (Romero et al., 2006) and (Romero et al., 2007). Several techniques could be used for personalization and recommendation such as classification, clustering, prediction, association rule, and sequential pattern. In (Tiffany et al., 2003), students were clustered with similar learning characteristics and applied collaborative filtering to provide paper recommendation.

A recommender agent using association rules has been used to recommend e-learning activities in (Zaiane, 2002). Providing to a student the next link or task to do within the adaptable educational hypermedia system AHA! was the aim of the recommender system described in (Romero et al., 2007). In this paper we are going to describe an automatic personalization approach for providing learning object recommendations for online students in e-learning systems. It is to be noted here that by adopting the term "Learning Object" we mean any digital educational resource used within the e-learning environment, it could be a course, a page, a simulation, i.e. all known formats of digital educational resources regardless their granularity. These learning objects are referenced within the e-learning system by their URL, they are also archived in log files or tracked in databases as URL references. Therefore, automatic recommendation of learning objects means generating a list of URLs referencing educational resources (hosted and/or created inside the e-learning platform) in order to guide and support the e-learners. The proposed approach is taking into account both the Web access history of learners as well as the content of the learning material. Web mining techniques in combination with an open source Web information retrieval system are used to enable an implementation that is not only open and scalable, but also fast to deploy. The recommender system we aim to develop should be considered as an external module or plugin that can be included easily in e-learning systems (courseware, LMS, etc) to give automatic personalization. This paper is arranged in the following way: first we describe the proposed approach and the corresponding phases of modeling and recommending. In Section 3, we present some implementations of the proposed methodologies. In Section 4, we make some experiments and evaluation. Finally, conclusions and future work are presented.

A framework for building automatic recommendations in e-learning platforms

Our proposed framework is composed of two modules: an off-line module which pre-processes data to build learner and content models, and an on-line module which uses these models on-the-fly to recognize student goals and predict a recommendation list. Recommended learning objects are obtained by using a range of recommendation strategies based mainly on content based filtering and collaborative filtering, each applied separately or in combination. The recommendation procedure is performed using the following tasks:

- **Preliminary offline mining of learners’ models** based on Web usage mining techniques. First, we gather web learners’ sessions and we apply a clustering approach to directly cluster these sessions. Each cluster contains similar sessions, showing similar interests of different learners. Each cluster can also be viewed as one learner’s model;

- **Preliminary offline mining of association rules** (e.g. “Resource A → Resource B”) from clustered sessions;

- **Preliminary offline crawling and indexing of learning resources**: this step consists of crawling the entire learning resources available in a course repository and forming an inverted index mapping each keyword to a set of pages in which it is contained;

- **Extracting user preferences from the learner’s active session** (set of URLs or list of terms extracted from these URLs);
• **Computing relevant links to recommend for the active learner** by applying a number of recommendation strategies.

The proposed approach, with main features depicted in (Figure 1), is essentially based on two components: the **Modeling phase** and the **Recommendation phase** (Khribi et al., 2008).

![Figure 1: Proposed personalization approach](image)

**Modeling phase**

*Learner Modeling*: Despite the availability of abundant educational resources and services, it is still difficult to decide which learning objects better match the student’s needs in a given situation, unless we accurately know the profile of that particular learner and his/her online behaviour. A learner’s model is composed by a range of relevant information about the student using the e-learning environment. Frequently common types of information used in learners’ models can include the learner knowledge, demographic information, preferences, learning styles, etc. These components are strongly connected to the application of the learner model. Generally, according to (Brusilovsky, 1994), two main categories are outlined: the domain specific information and the domain independent information. In our approach, we intend adopting a three components learner’s model: learner’s profile, learner’s knowledge, and learner’s educational preferences. Regarding the scope of this paper, we use solely the learner’s knowledge to represent the learner’s model. Other parts are being processed and are resumed in further papers. The learner’s model (reduced here to his/her learner’s knowledge component) can be represented by a sequence of weighted visited learning objects i.e. a vector of visited learning objects or curriculum elements in which the student was interested. Most approaches for user modeling heavily depend on user feedback. Users with common interests and levels can be grouped together. Feedback from one user can serve as a guide for information delivery to the other
users within the same group. The learner’s model is generally built based on a range of information, gathered through the user’s implicit and/or explicit feedback, describing the learner’s preferences on learning objects.

Basically, two main learner modeling approaches can be outlined: collaborative learner modeling and automatic learner modeling. The collaborative learner modeling approach requires students to provide explicit information about their preferences and needs. In the automatic learner modeling approach, gathering information is done rather automatically based on the online behavior and activities (i.e. implicit feedback) of students. Automated building of learner models involves the automated detection of all basic information composing the model i.e the learner’s knowledge in our case. Indeed, huge amount of data are collected continuously from the student interactions and browsing history, and capitalized automatically on server’s side and/or e-learning system database. To analyze these tracked data and make it fruitful to use we apply Web mining techniques. Data must be first pre-processed, using tasks including: data cleaning, user identification, and session identification. Corresponding to each student, we obtain set of sessions (i.e. several sessions gathered over a period of time). Let \( LO \) be a set of \( n \) unique visited learning objects: \( LO = \{ LO_1, LO_2, LO_3, \ldots, LO_n \} \), and let \( L \) be a set of \( m \) learners registered in a specific course within the e-learning environment, \( L = \{ L_1, L_2, \ldots, L_m \} \), the learner knowledge model \( LK_i \) corresponding to the learner \( L_i \in L \) is represented by a set of \( p \) sessions \( S_{ij} \) extracted from tracked data: \( LK_i = \{ S_{i1}, S_{i2}, S_{i3}, \ldots, S_{ip} \} \), where each \( S_{ij} \) is a subset of \( k \) weighted visited \( LO \):

\[
S_{ij} = \{ (LO_1, w(LO_1)), (LO_2, w(LO_2)), \ldots, (LO_k, w(LO_k)) \},
\]

where each \( LO_l \) is a weight associated with learning object reference \( L_{ij} \) in the session \( S_{ij} \) corresponding to the \( i^{th} \) learner \( L_i \). The knowledge model of a learner \( L_i \) (i.e learner model) can be represented by a matrix \( M(p, n) \) where \( p \) is the number of completed sessions and \( n \) the cardinality of unique visited learning objects:

\[
M(i, j) = \begin{pmatrix}
w(LO_1) & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & w(LO_n)
\end{pmatrix}
\]

**Group Modeling:** Once learners’ models are delimited properly, we apply a two-level model based collaborative filtering approach in order to organize the obtained models into groups of learners based on similarities and dissimilarities among their preferences. In the first level, we apply clustering techniques based on similarities and dissimilarities among preferred visited learning objects. A variety of clustering techniques can be used for clustering sessions. (Shahabi et al., 1997) described a prototype system that uses viewing time as the primary feature to describe a user session and then clusters the sessions using K-Means clustering. (Mobasher et al., 2000) used a multivariate K-Means algorithm to obtain transaction clusters and the Association Rule Hypergraph Partitioning (ARHP) technique to obtain usage clusters. (Yao et al., 2002) applied the Leader algorithm for clustering. In (Zhao et al., 2005), hierarchical clustering algorithm is presented and applied in CLUTO software for clustering high-dimensional datasets. Regardless of which method is used, the first clustering level will result in a set \( C = \{ C_1, C_2, \ldots, C_d \} \) of clusters where each cluster \( C_i \) a subset of student sessions representing a group of similar learners with similar access patterns. It should be noted here that the similarity measure used in clustering is the cosine similarity as shown in (1). In this equation, we have considered binary weight i.e. existence or non existence of a referred learning object in a session.

However, to give to the similarity measure more sense (semantic similarity measure), we can consider learning object content implicitly by taking into account a hierarchical web pages structure as in (2). In (3), \( S_u(i, j) \) is a URL to URL similarity function computed based on the amount of overlap between the paths \( P_i \) and \( P_j \) leading from the root of the website to any two URLs \( i \) and \( j \) (Nasraoui et al., 2008).

\[
S_{ij} = \begin{cases} 
1 & \text{if student accessed } j^{th} \text{ LO} \\
0 & \text{otherwise}
\end{cases}
\]

\[
S_{1,kl} = \frac{\sum_{i=1}^{N_u} \sum_{j=1}^{N_u} S_i^{(k)} S_j^{(l)} S_u(i, j)}{\sqrt{\sum_{i=1}^{N_u} S_i^{(k)^2} \sum_{i=1}^{N_u} S_i^{(l)^2}}}
\]

(1)

\[
S_u(i, j) = \min\left(1, \frac{|P_i \cap P_j|}{\max(1, \max|P_i, P_j|-1)}\right)
\]

(2)
The second level applied on each obtained cluster \( C_i \) (containing several sets of referred learning objects) consists of using, first, a frequent itemset mining algorithm that extracts frequently co-occurring referred learning objects in sessions belonging to each cluster. Then, association rules (AR) are extracted from these clusters. Association rules capture the relationships among learning objects references based on their co-occurrence across sessions. AR discovery methods such as the Apriori algorithm (Agrawal et al., 1994) can be applied. An association rule \( r \) is an expression of the form: \( A \Rightarrow B \), where \( A \) and \( B \) are itemsets, \( r \) must satisfy a minimum confidence threshold to be involved in recommended page set. The confidence \( \alpha_r \) of a rule \( r \) is given by \( \sigma(A \cup B)/ \sigma(A) \) and its support \( \sigma_r \) is the support of \( (A \cup B) \).

\[
\sigma(1_i) = \left| \left\{ S_j' \in C_j : I_i \subseteq S_j' \right\} \right| / |C_j|
\]

Thereby, further in recommendation phase, the new active learner is directly classified in one of the discovered groups of students (clusters), then personalized recommended links are provided by matching the learner current navigation with the AR of the corresponding group.

**Content Modeling:** Generally, content modeling involves applying indexing and text mining techniques (which are part of Web content mining). The originalities of our approach are twofold: (1) the use of the open source search engine Nutch (http://lucene.apache.org/nutch) in the content modeling phase, followed by content based filtering as a recommendation strategy, (2) The automated indexing of standard defined educational content thanks to the search engine’s powerful capabilities in the automated and scaled crawling and indexing phases.

To improve the educational content indexing phase covering two important levels: data (content) and meta-data (LOM), specific index fields used in LOM (Learning Object Metadata) are added to the native Nutch index. Thereby, accuracy of the final index, content search and recommendations are improved. Finally, we note that a byproduct of our crawling and indexing with Nutch is an available interface to use for explicit searches over the material if needed. Specific attributes to add into the Nutch index structure are given by the educational metadata providing descriptions and additional information (author, title, learning resource type, technical requirements, rights management, etc) about learning resources. These fields are added using a specific updated plugin. Corresponding information are loaded and inserted automatically to the inverted index thanks to XML files embedded to Standard defined Learning Objects and Nutch capability to crawl and parse XML files.

**Recommendation phase**

*Formulating a learner’s implicit query:* The learner’s implicit query is a set of referred learning objects recently visited by an active learner. Such query should be extracted implicitly from the recent navigation history of the learner and could be represented by a vector of referred learning objects or a vector of relevant terms describing these learning objects. This task is accomplished in two phases (1) delimiting the current active learner session, and (2) extracting URLs (Learning Objects references) of interest from this active session. Since the active user session is to extract from the Web log file (or tracked data in DB), we identify only the records representing the last \( W \) visited pages called the sliding window. This is done by taking into account the timestamp when a learner connected to the system and the timestamp of a recommendation demand. In order to express with more details learner preferences and interests, we can associate to each URL, composing learner sessions, a given weight. These weights can be binary, representing the existence or non-existence of an URL reference in the session; or they can be computed as a function of a number of features based essentially on the frequency of occurrence of URL within a session and the time a learner spends on a particular page as a manner to determine indirectly the fact that a learner liked or disliked the URL (Shahabi et al., 1997), (Yan et al., 1996). In addition to that, many other features could be added to compute page weight, but, generally, such features are considered as not a good indication of the user interest (Konstan et al., 1997). In the present work, we didn’t consider the URL weight. Let \( W \) be a fixed size for a sliding window, then if the active user session with \( W = 3 \) is \( \{A,B,C\} \), and the URL “D” was viewed last by a user, then the new sliding window becomes \( \{B, C, D\} \). It should be noted that \( W \) can lead to lower or higher recommendation coverage. In our case, we consider a fixed window size \( W = 3 \), so only the last three visited pages will affect the recommendation. Once we have obtained the sliding window pages, we translate it into a term vector. In fact, each page is mapped to a set of content terms characterizing it. We developed a java code (based on Nutch’s built in functionalities for parsing HTML pages, and a plug-in for stop word elimination) that returns, corresponding to a given page, the top \( K \).
frequent terms sorted by their frequencies. In our experiment (see section 3), we considered a fixed number of terms \( K = 3 \). Obtained terms are representing what we call a term vector (Figure 2).

**Figure 2: Term Vector Building process**

**Recommendation process:** The learner’s implicit query defined previously under both of its shapes constitutes the input of the recommendation phase. The recommendation process task is accomplished using basically: content-based filtering (CBF) and collaborative filtering (CF) approaches (Figure 3). First, we apply the (CBF) approach alone using the search functionalities of the Nutch search engine. We submit the term vector to the search engine in order to compute recommendation links. Results are ranked according to the cosine similarity of their content (vector of TF-IDF weighted terms) with the submitted term vector. Second, we apply the collaborative approach (CF) alone by comparing, first, the sliding window pages to clusters (groups of learners obtained in the offline phase by applying two-level model based collaborative filtering approach) in order to classify the active learner in one of the learner’s group. Then, we use the ARs of the corresponding group to give personalized recommendations. The current session window is matched against the “condition” or left side of each rule. (Example: “75.43 % of learners who visited http://cours.uvt.rnu.tn/rpl/cours/informatique/bases_donnees/base_donnee/chap6/menu.htm visited http://cours.uvt.rnu.tn/rpl/cours/informatique/bases_donnees/base_donnee/chap6/obje.htm and http://cours.uvt.rnu.tn/rpl/cours/informatique/bases_donnees/base_donnee/chap6/index6.htm”). All rules should be sorted by their confidence (highest to lowest) when forming the recommendations.

**Figure 3: Recommendation process**
It is worth noting that several recommendation strategies using these approaches have been investigated in our work. After applying a CF and CBF approaches alone, we included next the possibility to combine both of the recommendation approaches (CBF and CF) in order to improve the recommendation quality and generate the most relevant learning objects to learners. Hence, two approaches are to be considered: Hybrid content via profile based collaborative filtering with cascaded/feature augmentation combination, which performs collaborative recommendation followed by content recommendation (the reverse order could also be considered); and Hybrid content and profile based collaborative filtering with weighted combination, where the collaborative filtering and content based filtering recommendations are performed simultaneously, then the results of both techniques are combined together to produce a single recommendation set (Nasraoui et al., 2006).

In the Hybrid content via profile based collaborative filtering with cascaded/feature augmentation combination approach, we apply first CF approach giving as output a set of recommended links, then we apply CBF approach on these links. In fact, recommended links are mapped to a set of content terms in order to compose a term vector (top $k$ frequent terms), a parser tool must be used for this task. Finally, these terms are submitted to the search engine Nutch which returns the final recommended links.

In the Hybrid content and profile based collaborative filtering with weighted combination approach, the collaborative filtering and content based filtering are performed separately, then the results of both techniques are combined together to produce a single recommendation set. This process uses the following steps:

I. Step 1 is performed in the same way as in CF approach, the result is called Recommended Set 1;
II. Step 2 maps each LO references in the sliding window to a set of content terms (top $k$ frequent terms). Then these terms are submitted to the search engine which returns recommended links. This result is called Recommended Set 2;
III. Final collaborative and content based filtering recommendation combination: both recommended sets obtained previously are combined together to form a coherent list of related recommendation links, which are ranked based on their overlap ratio.
Experimentation and results

To implement and evaluate the proposed personalization approach, we conceived a system composed by a set of components, where each component is performing a number of services. The main features of the proposed recommender system are shown in Figure 4.

Component 0 extracts the learner active session from the Web log file, starting from the time that the learner connected to the platform until he/she asks for recommendations. Component 1 and Component 2 represent the input for the recommender system in two different forms. Component 1 extracts the W sliding window pages from the active learner session. The second input form is performed by Component 2 which extracts from Component 1’s results (sliding window) the top K relevant terms. Each page of the sliding window is transformed into a set of content terms. This task is performed using Nutch’s built-in functionalities for parsing HTML pages, and a plug-in for stop word elimination. Components 0, 1 and 2 are performed in an on-line mode. Component 3 concerns the phase of learner modeling which is done in an offline mode. We used the course repository of the Virtual University of Tunis, RPL platform (http://cours.uvt.rnu.tn) as an experimentation environment for our system. Web usage logs are collected from the RPL log files (Apache web access log files http://httpd.apache.org/docs/2.0/logs.html) in the period between February and July 2007. RPL Log files were rotated based on daily directives, thus generating 180 large log files with a total of 3,049,986 requests. The log format was ECLF (Extended Common Log Format), further enriched with added user authentication information to make the session extraction error free (using an embedded session Id mechanism, added via the Apache configuration and our RPL code). Since the collected log data contain many uninteresting elements (graphics, icons, requests generated by crawlers/bots, etc), it must first be pre-processed. Starting with 3,049,986 requests, the cleaning operations resulted in only 594,325 nontrivial requests. Figure 5 shows the variation of request numbers before and after data cleansing per month. After cleaning the original logs, we sessionized the remaining requests. Figure 6 shows the variation of the number of resulting unique usage sessions per month.

We complete all missing SessionID attributes in the database tables and we used mainly the algorithm below to process this task, then, we determined the number of unique usage sessions using simple SQL queries.

```
Begin
For each row L_i of T
If L_i[SessionID] is empty
Add_Session(L_i, L_i)
End

Add_Session(L_i, L)
Begin
If (L_i[TimeStamp]-L[TimeStamp]) < 30
L_i[SessionID] = L_i[SessionID]
Else
Add_Session(L_i, L_i)
Else
L_i[SessionID] = New_SessionID()
End
End
```

![Figure 5: Pre-processed RPL logs per month](image)

![Figure 6: Variation of unique sessions per month](image)
If \((L_i[\text{Referrer}] = L_i[\text{Url}] \text{ And} \n L_i[\text{Agent}] = L_i[\text{Agent}])\)

Once we have the session file, we applied, first, CLUTO soft (http://glaros.dtc.umn.edu/gkhome/views/cluto) using a hierarchical partitional clustering algorithm (with cosim similarity and \(K=10\) clusters) to cluster learners’ sessions. Then, each cluster is represented by a vector composed by pages with highest membership (occurrence). In the second level, association rules are extracted from each cluster, we used Goethals apriori algorithm implementation to compute frequent items, and Goethals association rules implementation to mine AR within each cluster (http://www.adrem.ua.ac.be/~goethals). Component 4 concerns the phase of content modeling. To experiment Nutch crawling process, we built an URL file containing 73469 URLs representing all learning objects available in RPL repository. The Nutch crawler was invoked using the following command line: (nutch crawl urls -dir crawldir -depth 6). The crawler uses the URL file for fetching, parsing and indexing the URLs, thus creating an inverted index thanks to XML files describing files the Standard Defined Learning Objects and Nutch capability to crawl and parse XML files. In component 5, R1 et R2 represent the two main recommendation strategies used to compute what to recommend to the learners. Each recommendation strategy uses as input results returned by Components 1 and 3 and/or Components 2 and 4. Component 5 finally performs the task of delivering recommended links by combining the use of various recommendation approaches or by using them separately based on guidelines given by Component 6 which represents the configuration module specifying a set of entry details to the whole recommender system, such as recommendation strategy and the variation of related parameters (\(K, W, N, \) etc). For example, concerning the two-level collaborative filtering approach, to make recommendation, Component 5 starts by classifying the learner’s active session under the shape of vector of visited learning objects to the closest cluster (based on Component 1 and 3) using the KNN algorithm (with cosim similarity and \(k=1\)). Then, matching task between learner’s active session and antecedent of AR within the selected cluster is accomplished to produce finally the set of recommended links. Figure 7 shows a learner’s activities in RPL platform when browsing courses. Figure 8 shows the recommendations based on Content based Filtering using Nutch. The bottom window contains a list of set-link (shown as “Ensemble i”), each set-link is used to formulate a query whose top search results, as a set, are accessible by clicking the corresponding set-link. These results are shown in the window above. The top-most window contains the actual content of a recommended page. Figure 9 shows a list of recommended links based on Collaborative Filtering approach. Notice that for CF recommendations, we also added explanations in the form of the association rules that led to matching inputs. The bottom window of Figure 10 shows list of set-links pointing to a list of recommended links based on cascaded hybrid (collaborative filtering followed by content-based filtering) as shown in the window above. The top-most window contains the actual content of a recommended page.

**Evaluation**

In our experiments, we used the data file obtained after data preprocessing and sessionization, this file contains 11542 unique sessions. We select about 70% of these sessions to compose the training set, the remaining sessions are used for evaluation. Each session \(Si\) in the evaluation session file is divided into two parts. The first \(n\) \(LO\) references in \(Si\) are used as the sliding window (the learner’s query) to generate recommendations. The remaining portion of \(Si\) is used to evaluate the generated recommendations. We considered that the maximum size of the sliding window \(|w|\) is 4, \(|w|<=4\). Let \(Si(w)\) be a set of learning objects composing the active session used by the recommender engine to produce a recommendation set, \(t\) is a recommendation threshold. The recommender engine takes as inputs \(Si(w)\) and \(t\), \(R(Si(w), t)\) represents a set of recommended learning objects which score is at least \(t\). Let \(ESi\) a set of remaining learning objects references in a session \(Si\) within the evaluation session file. The experimental evaluation is based on two metrics: precision and recall defined as following: Precision\((R(Si(w), t)= R(Si(w), t) \cap ESi) / | R(Si(w), t) \cap ESi|\) and Recall\((R(Si(w), t))= R(Si(w), t) \cap ESi) / | ESi|\). Precision measure represents the proportion of relevant recommendations to the total number of obtained recommendations, whereas Recall measure or coverage represents the proportion of relevant recommendations to all learning objects that should be recommended (Mobasher et al., 2001). In our experimentation we measured both precision and recall and we varied.
recommendation threshold from 0.4 to 1.0. We considered only the recommendation process based on the two-level collaborative approach.

Figure 7: The active learner is browsing material (active user session)

Figure 8: Recommendation using content based filtering approach

Figure 9: Recommendation (blue links) using collaborative based filtering approach. (Explanations based on collective association rules are listed at the bottom)

Figure 10: Recommendation using cascaded hybrid (collaborative filtering followed by content-based filtering)
First, we applied CLUTO soft using a hierarchical partitional clustering algorithm (with cosim similarity and $K=10$ clusters) to cluster learners’ sessions in the training set. Then, each cluster is represented by a vector composed by pages with highest membership (occurrence). In the second level, association rules are extracted from each cluster, we used Goethals apriori algorithm implementation to compute frequent items, and Goethals association rules implementation to mine AR within each cluster, we considered Confidence=0.4, Support=0.04. In the recommendation phase, we start by classifying the active session $S_i(w)$ in the closest cluster using the $KNN$ algorithm (with cosim similarity and $k=1$). Then, matching task between active session and antecedent of AR within the selected cluster is accomplished to produce finally the set of recommended links. We performed all experiments varying window size from 1 to 4. Finally, for a given recommendation threshold $t$, the mean over all sessions in the evaluation file was computed as the overall evaluation score for each measure. Curves below (Figure 11 and Figure 12) show that precision values increase with the increase of sliding window size and decrease when the number of recommended links increase (conversely for Recall).

For each session $S_i$ in the Evaluation_session_file $i$ in $\{1..p\}$

- Select the active session $S_i(w)$ for recommendation and the Remaining lo set $ES_i$ for evaluation
- Select the closest cluster $C_j$ for $S_i(w)$
- For each Rulek in AR file
  - If(matching_rulek($S_i(w)$,rulek(cond)))
    - Candidate_lo_Si(lo,rank_lo]+=rulek(conseq, rulek(confidence))
    - FinalCandidate_lo_Si(lo, rank_lo) = Unique_Candidate_lo_Si(lo, final_rank_lo)
      with final_rank_lo=sum(over all rank_lo for the same lo)
- For recommendation threshold $t$ varying from 0.4 to 1.0
  - If final_rank_lo > $t$ lo is added Recommended_Set(lo)
    - Precision_t+=((Recommended_Set(lo)\ESi)/Recommended_Set(lo)
    - Recall_t+=((Recommended_Set(lo)\ESi)/ESi
- For recommendation threshold $t$ varying from 0.4 to 1.0
  - Precision_Tot_t=Precision_t/p
  - Recall_Tot_t=Recall_t/p

**Figure 11**: Precision of Recommendations  
**Figure 12**: Recall of Recommendations

**Conclusions and future work**

In this paper, we have outlined the general principles of a new approach to perform personalization in e-learning platforms by resorting to a recommender system relying on web mining techniques and scalable search engine
technology to take care of one of the crucial steps in personalization, which occurs in the "online" phase to compute the recommendations against a possibly massive repository of educational resources in "real time". In the modeling phase, we used Nutch’s automated crawling and indexing techniques as well as standardized educational content metadata to build content models, and Web usage mining techniques (clustering and association rule mining) to build learner profiles. Hybrid recommendations (based on CBF and CF) were used in the recommendation phase. We are currently exploring several techniques and strategies in the modeling and recommendation phase in more detail, and performing more evaluations. We are also studying the possibility of integrating educational preferences in the learner’s model such as learning styles, media types, etc. The learner’s model to consider, in the future work, should be composed of three main components: learner’s profile, learner’s knowledge and learner’s educational preferences. All these components should be detected automatically within e-learning systems. After construction of the student models, we build group models using a three-level collaborative modeling approach. We expect this enrichment of the learner’s model to increase the quality of learning object recommendations especially from an instructional point of view.

Acknowledgments

This research was partially supported by National Science Foundation CAREER Award IIS-0133948 to O. Nasraoui.

References


A Framework for Semantic Group Formation in Education

Asma Ounnas, Hugh C Davis and David E Millard

Learning Societies Lab, School of Electronics and Computer Science, University of Southampton, Highfield, Southampton, Hampshire, SO17 1BJ, UK // ao05r@ecs.soton.ac.uk // hcd@ecs.soton.ac.uk // dem@ecs.soton.ac.uk

ABSTRACT

Collaboration has long been considered an effective approach to learning. However, forming optimal groups can be a time consuming and complex task. Different approaches have been developed to assist teachers allocate students to groups based on a set of constraints. However, existing tools often fail to assign some students to groups creating a problem well known as “orphan students”. In this paper we propose a framework for learner group formation, based upon satisfying the constraints of the person forming the groups by reasoning over semantic data about the potential participants. The use of both Semantic Web technologies and Logic programming proved to increase the satisfaction of the constraints and overcome the orphans’ problem.

Keywords

Group Formation, E-Learning, Constraint Satisfaction, Collaborative Learning, Teams.

Introduction

Many approaches to learning and teaching rely upon students working in groups. Research in many disciplines has shown that learning within groups improves the students’ learning experience by enabling peers to learn from each other. To form groups, students can be either allocated to groups randomly, self-select each other, or be appointed to a group by the teacher based on some criteria related to the collaboration goals. These criteria are usually expressed as a set of conditions, typically referred to as constraints, such as restricting the groups to be mixed in gender or skills (Ounnas, 2007b).

For the teacher, forming groups manually can be both difficult and time consuming. For this, researchers have been investigating several techniques for automating this process through the use of computer-supported group formation (CSGF). Similar to manual group formation, the challenges of CSGF lie in modeling the students’ data, the teacher’s constraints; and negotiating the allocation of students to groups to satisfy these constraints. However, existing tools often fail in allocating all students to groups, leaving some students unassigned to any group after the formation (Redmond, 2001), (Tobar, 2007). This problem is usually referred to as the orphan students problem.

In previous work (Ounnas, 2007b), we discussed some of the existing CSGF techniques in terms of the constraints and the selection of members in the groups. We also discussed the potential of using Semantic Web technologies (Berners-Lee et al., 2001) in providing meanings to the students’ descriptions and constraints. In this paper, we propose a framework that is capable of efficiently automating the formation of students’ groups by reasoning over the students’ semantic data and the list of constraints specified by the teacher. We use the efficiency of both Semantic Web technologies and logic programming in modeling the problem of group formation as a constraint satisfaction problem (Kumar, 1992). The next section of the paper describes our motivation behind the research based on existing literature and the results obtained from a case study. We then describe the structure of the proposed framework and explain its components. Afterward, we discuss the evaluation of the proposed framework. Finally, we describe our future work for improving the performance of the framework and discuss some of the relevant issues with its evaluation.

Motivation

In order to understand the issues rising with forming efficient groups of learners, we analyzed the existing applications and efforts to automate this process. We identified the limitations of the existing tools and the need to design a framework that enables the delivery of well-formed groups based on constraints from a multidimensional space. To understand the nature of the possible constraints we carried an observational study with a class of undergraduate students at the University of Southampton.
Existing applications

In (Hoppe, 1995) Hoppe introduced an intelligent tutoring system that allows the learners to initiate a group formation when they have a problem (a learner-helper group). Based on the learners’ models, the system displays a list of all potential peer learners that can help; the learner then selects a helper from the list, and the latter can accept or reject the invitation to help the learner. Parameters here are based on learning experience and competency criteria in the subject of the collaboration. In Mühlenbrock (2005) and (2006), context information such as the learner’s geographical location from PCs, Phones, and PDAs were added to the model. Unfortunately, no evaluation of the application was provided by the authors.

A team from Osaka University in Japan (Ikeda et al., 1997) and (Inaba et al., 2000) introduce Opportunistic Group Formation (OGF) where an intelligent system detects the appropriate situation to start a collaborative learning session and sets up a learning goal for the learner. The system takes into account the modeling of learning goals for each learner. Based on individual goals as well as the whole group, the system negotiates with the agents of all the learners in order to come to an agreement and to form a learning group so that each member of the group can obtain some educational benefit. Unfortunately, there is no literature on the architecture of the developed systems or their evaluation.

In similar research (Soh et al., 2006), (Zhang, 2005), the authors introduce a multiagent intelligent system called I-MINDS where the instructor, each student and each group is represented by an intelligent agent. The student agent profiles the student and finds compatible students to form the student’s “body group”. The agents communicate, and form coalitions dynamically. For the group formation, each student agent bids to join its favorite group based on their previous performance in group activities. Therefore, the formation is constrained by the learner’s previous performance. The collaboration, formation, and learner profiles updates are all processed in real-time. The student’s profile in this research is built dynamically based on how active is the student during the real-time collaboration. At the beginning of this research (Soh, 2004), the author intended to provide a group formation based on positive interdependence and hence joint intentions where the students depend on each other for goal satisfaction, rewards, resources, division of labor, roles, and so on. However, when I-MINDS was introduced (Soh et al., 2006), the authors did not mention how this theory was put into practice, and only mentioned that their application considers the students’ previous performance as a constraint to the formation. Soh et al. evaluated their IMINDS system by measuring its effectiveness in terms of its ease of usability by the instructor and the students. The group formation itself was evaluated against the performance of the teams, measured based on the teams’ outcomes and their responses to a series of questionnaires that evaluates team-based efficacy, peer rating, and individual evaluation.

Also supporting Opportunistic Group Formation systems, in (Wessner and Pfister, 2001), the course author defines at which points in a distributed web based course a collaborative activity should occur. The system then uses knowledge about the collaboration context in real-time such as whether the student has performed this collaboration before, how often, and how fast in order to form appropriate groups. The formation here follows a self-selected approach. Although the authors did not present any results of evaluating their system, they mentioned that the comprehensibility of the group formation algorithms and the satisfaction of learning groups to be a key factor of the overall approach acceptance.

On recommending expert collaborators, Vivacqua and Lieberman (2000) introduces Expert Finder, an agent that automatically generates user models by classifying both novice and expert knowledge of the participants. The agent autonomously analyzes documents created in the course of routine work to rank the experts for recommendation to the learner who initiated the expertise search. For evaluation, the authors compare the results (the formed groups) generated from the system to manual generated results of the same participants’ sample. This technique is frequently used in recommender systems (McDonald, 2001).

Redmond (Redmond, 2001) introduces a computer program to aid the assignment of students projects groups. This technique is used to form instructor-based group formation for all (part time and distance) learners in the class simultaneously. The students are grouped based on the time slot they prefer to do the group work in, and then allocate the projects to the groups based on the members’ preferences in the group. The group formation is processed using a greedy algorithm where the program starts with the tightest constraint – the student with the fewest time slots rated highly – and tries to find a compatible group for them. This process repeats until all students have been assigned to a group. The formed groups are manually checked for even distribution of grades, and the students who
are left unassigned are manually allocated to groups. To measure the efficiency of the program, the author introduced an evaluation formula that calculated the rating of group assignments by subtracting an unassigned penalty representing the program failure in assigned some students' performance from the sum of all formed group overall rating.

In Christodoulopoulos (2007), the authors presented a web-based group formation tool that supports the instructor to automatically create homogeneous and heterogeneous groups based on up to three criteria and the learner to negotiate the grouping. The tool employs a clustering algorithm (Fuzzy C-means) for homogeneous grouping while heterogeneous groups are generated using Random Selection algorithm. For each student, the clustering algorithm gives the probability of the student belonging to each group. This helps the instructor to manually adjust the formation since the generated clusters may not be of same size. The probabilities enable swapping students who are unsatisfied with their allocation. The preliminary evaluation of the tool was satisfactory although it was tested on only 18 groups with one criterion (constraint on Learning Styles).

In Tobar et al (2007), the authors introduced a rule-based tool that aims at reducing the time teachers spend creating groups for learning. The tool takes into account the students’ characteristics that are required by the rule (hard constraint that can not be violated). The characteristics available in this tool are taken from the IMS learner specification (Wilson, 2002). The results returned from the tool can be manually modified by the instructor. Unfortunately, there was no evaluation of the performance of this tool.

In Lugano et al (2004), the authors studied data from self-rated questionnaire together with statistics of the learners’ real activity in a collaborative learning environment called EDUCOSM. The authors considered the students’ motivation (learning goal orientation) and social skills (social group roles). The results of analyzing self-perception in actual behavior showed a low correlation of pre-test results with learning outcomes. The authors explained this observation by the high initial expectations being lowered later by factors such as high workload or technical problems with the system.

In de Faria et al (2006), the authors introduced an approach of forming groups for collaborative learning of computer programming. The groups were formed based on the students’ programming style generated by a tool implemented to automatically assess the style of the programs submitted by the students. Analyzing the students’ programs assists in finding characteristics that evidence significant differences such as program quality, which would be relevant enough to motivate the students to discuss them.

In Graf et al (2006), the authors propose a mathematical model for building heterogeneous groups based on the students’ personality traits (group work attitude, interest for the subject, achievement motivation, self confidence, and shyness), their level of performance in the subject, and fluency in the language of instruction, where each of these attributes is ranked on a one to three scale. The authors use the Ant Colony Optimization algorithm to allocate each student to the most appropriate group that would maximize the diversity of that group while keeping the deviation between the groups minimum. The authors show that their approach is scalable (around 500 student) despite the problem being NP-hard.

In Cavanaugh et al (2004), the authors describe Team-Maker, a web-based system that aims at reducing instructors’ time in allocating students to groups. The system takes some students characteristics such as gender, skills, and students’ schedules, and the instructor’s criteria for the creating of homogeneous or heterogeneous groups, and applies a Hill climbing algorithm to get the optimal solution. The authors show that the system outperforms manual group formation, but does not mention the complexity of the system or how good it performs as the number of constraints (instructor’s criteria) grows.

In Wang (2007), the authors introduce a computer-supported heterogeneous grouping system called DIANA. The system uses a genetic algorithm to form fair groups in terms of heterogeneous grouping such that all groups have the same level of diversity. The system uses the students’ characteristics (thinking styles) collected from questionnaires. It takes up to 7 variables and allocates 3 to 7 members per group. The evaluation of the research on a class of 66 students showed that DIANA performs better than random allocations to groups. Although, the authors did not discuss the complexity or scalability associated with the application of the algorithm. Table 1 shows a summary comparison of the discussed Computer Supported Group Formation applications. We refer to the applications that are not provided with a name with their first author name.
Table 1. Existing CSGF applications in e-learning

<table>
<thead>
<tr>
<th>Formation features</th>
<th>Approach</th>
<th>Principle</th>
<th>Algorithm</th>
<th>Modeled students characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-selecting</td>
<td>Instructor Based</td>
<td>Opportunistic</td>
<td>Modeled students for all students in the class</td>
</tr>
<tr>
<td>Hoppe</td>
<td>✔</td>
<td>✔</td>
<td>×</td>
<td>Rule/inference based</td>
</tr>
<tr>
<td>Inaba</td>
<td>✔</td>
<td>✔</td>
<td>×</td>
<td>Multi-agent System</td>
</tr>
<tr>
<td>Soh</td>
<td>✔</td>
<td>✔</td>
<td>×</td>
<td>Multi-agent System</td>
</tr>
<tr>
<td>Wessner</td>
<td>×</td>
<td>✔</td>
<td>✔</td>
<td>Multi-agent System</td>
</tr>
<tr>
<td>Vivacqua</td>
<td>✔</td>
<td>×</td>
<td>×</td>
<td>Profile Matching</td>
</tr>
<tr>
<td>Redmond</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>Greedy algorithm</td>
</tr>
<tr>
<td>DIANA</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>Genetic algorithm</td>
</tr>
<tr>
<td>Team-Maker</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>Hill Climbing</td>
</tr>
<tr>
<td>Graf</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>Ant Colony optimization</td>
</tr>
<tr>
<td>Tobar</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>Rule based</td>
</tr>
<tr>
<td>Christodoulopoulos</td>
<td>×</td>
<td>✔</td>
<td>×</td>
<td>✔</td>
</tr>
</tbody>
</table>

Limitations of existing applications:

From the literature, we see that in terms of constrained group formation complexity (Ounnas, 2007):

Modeling

- Most systems only model a fixed set of parameters, which does not allow for the formation of different types of groups, and hence the implementation of different collaborative activities (only supports some types of teams).
- None of the existing efforts discuss the performance of the relative application in handling the group formation when the data about the user is incomplete, for example, if a new student with no record in the university joins the collaboration activity.

Constraint satisfaction

- Many systems use Opportunistic Group Formation (Wessner and Pfister, 2001), (Hoppe, 1995), (Soh et al., 2006), (Inaba et al., 2000), which does ensure satisfaction of the participants in the group through negotiation, but does not discuss the efficiency of the negotiation if all students in the class are grouped simultaneously. In addition to that OGF is usually more beneficial in short-term groups. In addition to this, these systems are based on self-selecting group formation (Hoppe, 1995), (Inaba et al., 2000), (Vivacqua and Lieberman, 2000), (Soh et al., 2006), which is not the most efficient approach in forming teams for learning, as it does not ensure balanced grouping.
As observed in using existing group formation tools, another common problem in forming groups is “the orphans problem”; these are the students who remain unassigned to any group at the end of the formation. In existing tools, such as (Redmond, 2001) and (Tobar, 2007), this problem remains unsolved. Instead, most tools return the names of the orphans for the instructor to allocate them manually to some group, or rearrange the formation by swapping the orphans with other members, the fact that decreases the efficiency of the automated formation.

Based on the reported results, most applications can only take a small fixed number of constraints. So far, DIANA seems to handle the highest number of constraints, which is currently limited to 7, and only for homogeneous grouping. We hypothesize that is fact is related to the limitation and complexity of the algorithms implemented in the tools and therefore rises issues on scalability of the systems.

Evaluation

In addition to the luck of providing results on the performance of the applications in some of the literature, a limitation of most group formation applications is the exclusive reliance on the groups’ performance measures indicators such as members’ responses to questionnaires or post-tests to draw inference about the group formation system performance. From a learning viewpoint at least, group formation efficiency is clearly a multi-dimensional concept, which implies that multiple efficiency indicators besides perceived performance need to be employed. While different formation constraints might result in different formulas for calculating efficiency, these constraints can be related to group formation efficiency in a more abstract way. If so, consideration of defining this relation together with other group formation related measures is required.

Observational study

To match the growing need of forming groups with higher flexibility, we started analyzing what constraints do teachers consider when forming groups. We studied the possible students’ features that can be relevant to forming different types of groups by investigating the available literature on collaborative learning theories (Ounnas, 2007b), and asking teachers what constraints they employ for different educational goals.

As a case study on group formation, we conducted an observational study with 67 undergraduate students taking a software engineering group projects course (SEG) in the School of Electronics and Computer Science at the University of Southampton. The students were manually grouped by the course organizers into 11 groups of 5 to 6 students, based on the following constraints:

- All groups have to be balanced in terms of the students’ previous grades to ensure that all groups have an equal opportunity in performing well in the project.
- To avoid minorities, a female cannot be allocated to an all-male group to prevent her from being cast away by the members.
- International students from the same country can’t be all members of the same group.

The module organizers used a script to allocate the students based on their marks, then manually swapped some of them to redistribute females and international students. To analyze the dynamics of the groups and how other criteria affect them, we distributed two questionnaires to the class:

**Questionnaire (1):** at the beginning of the course, we asked the students to fill in a form to get information about their previous experience in software engineering, teamwork, their gender, nationality (to detect minorities), and Belbin team roles to check which role can each student play within their group. Belbin roles are typically used in industry and training activities to discover the best roles a participant can play in a group (Belbin, 2004). There are 8 Belbin roles, and according to these roles; a balanced team is composed of:

- One leader: *Coordinator (CO)* or *Shaper (SH)*, and not both in the same group to avoid conflicts,
- A *Plant (PL)*: to stimulate ideas and insure creativity,
- A *Monitor/Evaluator (ME)* to maintain honesty,
- One or more *Implementers (IM)* to executed actions, *Team Worker (TW)* to ensure cooperation in the group, *Resource Investigator (RI)* to explore opportunities and secure resources, or a *Completer/Finisher (CF)* to ensure all tasks are completed on time.
Each person usually plays more than one Belbin role within the team. However, a member usually scores high in only one or two roles. In our study, we collected both the first and second roles for each student.

Due to some students dropping out of the course and others not filling in the questionnaires, we collected data from 9 groups out of the whole 11. Table 2 illustrates the results collected from questionnaire (1) showing Belbin roles in each group. The numbers in the cells demonstrate how many members in the group have that role as their strongest role, where the distribution of the latter is the average of the first and second strongest roles.

Table 2. Results of observational study (distribution of Belbin Roles)

<table>
<thead>
<tr>
<th>Group</th>
<th>IM</th>
<th>CO</th>
<th>SH</th>
<th>PL</th>
<th>RI</th>
<th>ME</th>
<th>TW</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>

Questionnaire (2): at the end of the course, we distributed a 17 questions based evaluation form where the student is asked to rank the key elements that measure their group performance, dynamics, and the individual satisfaction with the group work on a 1 to 6 scale. In particular, we analyzed creativity, motivation, leadership, group cohesion, satisfaction with contribution of members and the group output.

Given that in some groups, only one or two students returned the questionnaire, we were only able to use the data from groups 1, 3, 4, 5, 7, and 8. Table 2 shows these groups in shaded color. The results showed that the majority of the groups were satisfied with the group output (the software), and no members (minorities) were isolated which can be related to the fact that the teams were formed to be balanced in terms of grades and gender. However, constant conflicts were reported in the groups that had no leader or more than one strong leader (groups 1, 7 and 8). The groups that did not have a plant member such as groups 3 reported a lack of innovation, while groups with a Plant responded well (group 1 and 5).

From the study, we observed the relation and effect of possible group formation constraints on the students’ perceived satisfaction. However, despite the benefits of having a number of constraints in achieving the educational goal of the collaboration, negotiating the students’ allocation to groups manually gets more complex and time consuming as the number of constraints grows, even if the teacher had the required data about the student.

The case study provided us with an initial understanding of the domain characteristics and relevant problems in forming groups, which support our findings from analyzing existing literature in the area.

Our analyses from both the literature and the case study yielded various ideas for possible computer support in both modeling the constraints and evaluating the formation of groups.

Semantic group formation framework

To overcome the complexity of allocating students to groups, we propose a framework to assist the teacher in forming groups based on their chosen set of constraints. The framework handles the group formation process based on the following concepts:

- Modeling the students’ features: we model a large range of features that can be considered for different group formations using the concept of Semantic Web ontologies (Berners-Lee et al., 2001), which can form a reliable
dynamic learner profile (Ounnas, 2007b). In this context, semantic modeling provides meaningful descriptions of the students and the relationships between them.

- Negotiating the group formation: we express the students’ allocation problem as a Constraint Satisfaction Problem (CSP) (Kumar, 1992). The negotiation process of allocating each student to their most appropriate group can then be handled by a constraint satisfaction solver.

We emphasize that, in this research, we are not concerned with proving that any particular set of constraints leads to better results in terms of the performance of the groups; neither do we claim that any particular algorithm leads to best grouping. Figure 1 shows an overview structure of the framework, which is based on the following components:

![Figure 1. Semantic group formation framework](image)

**The Student Interface**

The student can enter their data through a web-based form composed of four parts: the student’s personal data, a list of their friends, their interests and preferences, information about their course such as the modules they are taking, and a list of their friends/colleagues taking similar courses. The students can update their data at any time. To avoid the form style in the interface, and to make it more desirable to students to update regularly, we aim to make it a Web 2.0 like interface such as the ones used by social networking sites.

**The Ontology**

We created an ontology called Semantic Learner Profile (referred to in this paper as SLP). The ontology is available at [http://users.ecs.soton.ac.uk/ao05r/slp.owl](http://users.ecs.soton.ac.uk/ao05r/slp.owl). The ontology extends the Friend Of A Friend ontology (FOAF, available at [http://xmlns.com/foaf/spec/](http://xmlns.com/foaf/spec/)), an existing ontology that describes people for building communities and social groupings. The learner’s characteristics that the ontology describes were chosen based on a comparison of existing learner profiles such as PAPI, IMS LIP and eduPerson (Ounnas, 2006). Therefore, the ontology describes a large range of student’s personal, social, and academic data such as learning styles, preferred modules, topics, and collaborators (Ounnas, 2007a). The semantic representation of these data, to which the instructor constraints can be mapped to, allows inferences to generate more data. This feature of using semantics enables the framework to handle incomplete data in a more effective way (this is explained in more details in section 4). Since the FOAF ontology is currently very popular (Ding et al., 2005), employing it would allow using data from any other ontology that can be mapped to FOAF.
Once the student submits the profile data through the student interface, an RDF file is created (FOAF + SLP). The file is processed using Jena, a Semantic Web inference engine (Carroll et al., 2004), and instances of the ontology are then stored in a database. Figure 2 shows an example of a student’s FOAF file extended with the SLP ontology. In this figure, the file holds information about the student’s name, gender, Belbin role, preferred module, topics of interest, and friends (classmate).

```xml
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:foaf="http://xmlns.com/foaf/0.1/"
   xmlns:slp="http://www.ecs.soton.ac.uk/~ao05r/slp.owl">
   <foaf:Person rdf:nodeID="asma">
      <foaf:name>Asma Ounnas</foaf:name>
      <foaf:gender>Female</foaf:gender>
      <slp:belbin>Implementer</slp:belbin>
      <slp:interest>e-Learning</slp:interest>
      <slp:interest>Semantic Web</slp:interest>
      <slp:preferredModule>CS1004</slp:preferredModule>
      <slp:classmateOf><foaf:name>Ilaria Liccardi</foaf:name></slp:classmateOf>
   </foaf:Person>
</rdf:RDF>
```

*Figure 2. An example student FOAF+SLP profile*

---

Welcome to GroupMe: The Semantic Group Formation Tool

Generate a New Formation

Please select the group type:

- **Teams**: Number of members per group: 5

Please select the formation constraints:

- **Constraint 1**: Interests: Homogeneous, Weight: 1, Priority: 1
- **Constraint 2**: Learning Style: Heterogeneous, Weight: 4, Priority: 4
- **Constraint 3**: Gender: Not Equal One, Weight: 3, Priority: 3
- **Constraint 4**: Belbin: Functional, Weight: 2, Priority: 2

*Figure 3. An example student FOAF+SLP profile*
The Instructor Interface

Through this web-based interface, the instructor can select which constraints they care about for the formation they are initiating such as the student’s gender, their team role, their learning style and so on. So far, it is possible to add constraints on the following characteristics: gender, nationality, age, previous marks (for competency), team roles, interests, and learning styles. Other characteristics such as preferences and trust values between participants will be implemented soon.

For each constraint, the teacher can select the condition on this characteristic, for example the groups are to be heterogeneous in gender and mark distribution, homogenous in learning style, and so on. They are also provided with an option that enables them to set a priority value for each constraint. Ranking the importance of the constraints to the group formation enables the application to manage compromises based on these priorities as explained in the next section.

The group generator

As the core component of the framework, the group generator is responsible for negotiating the allocation of students into groups. The generator is based on a DLV solver, an implementation of disjunctive logic programming, used for knowledge representation and reasoning. DLV’s native language is Disjunctive Datalog extended with constraints, true negation and queries (Leone et al., 2006), where Datalog is a query and rule language for deductive databases that syntactically is a subset of Prolog.

DLV performs a simple forward checking algorithm (Kumar, 1992) on the data provided by the learners and the instructor in order to allocate students to groups. The students’ data is automatically transformed from the SQL database to an Extensional Database (EDB) in the form of predicates that the solver can read as an input. Figure 4 shows an example of this knowledge base where predicates of the form “student(name, role, gender)” show the student’s family name, Belbin role, and gender.

---

**Figure 4. Example DLV program**

---

<table>
<thead>
<tr>
<th>Example input from the knowledge</th>
<th>DLV Constraint Satisfaction Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>student(green,shaper,female).</td>
<td>%Number of student members per group as specified by the instructor nmembers(5).</td>
</tr>
<tr>
<td>student(jensen,shaper,male).</td>
<td>% Guess if participant P is a member of group G or not.</td>
</tr>
<tr>
<td>student(patterson,implementer,male).</td>
<td>member(N,B,S,G) v not member(N,B,S,G) : student(N,B,S), group(G).</td>
</tr>
<tr>
<td>student(jones,evaluator,male).</td>
<td>% Each student should be a member of only one group</td>
</tr>
<tr>
<td>student[harris,plant,male].</td>
<td>:- student(N,...), not #count(G: member(N,...,G)) = 1.</td>
</tr>
<tr>
<td>student(baker,plant,male).</td>
<td>% Each group should have C members or less (in this case C = 5)</td>
</tr>
<tr>
<td>student(watson,worker,female).</td>
<td>:- group(G), nmembers(C), not #count(N: member(N,...,G)) = C. [1:4]</td>
</tr>
<tr>
<td>student(williams,coordinator,male).</td>
<td>% Number of females per group cannot be equal to exactly 1</td>
</tr>
<tr>
<td></td>
<td>:- group(G), #count(N: member(N,...,female,G)) = 1.</td>
</tr>
<tr>
<td></td>
<td>% Distribute Belbin roles such that every group has one leader</td>
</tr>
<tr>
<td></td>
<td>(shaper or coordinator, and one plant, and at least one implementer)</td>
</tr>
<tr>
<td></td>
<td>:- group(G), not #count(N: member(N,shaper,...)) &lt;= 1. [1:3]</td>
</tr>
<tr>
<td></td>
<td>:- group(G), not #count(N: member(N,coordinator,...)) &lt;= 1. [1:3]</td>
</tr>
<tr>
<td></td>
<td>:- group(G), member(N,coordinator,...), member(M,shaper,...), N != M. [1:2]</td>
</tr>
<tr>
<td></td>
<td>:- group(G), not #count(N: member(N,plant,...)) &lt;= 1. [1:3]</td>
</tr>
<tr>
<td></td>
<td>:- group(G), not #count(N: member(N,implementer,...)) &lt;= 1. [1:1]</td>
</tr>
<tr>
<td></td>
<td>% Model confidence showing constraints with priority 2 as violated.</td>
</tr>
<tr>
<td>Cost=10:1.1:1.21:10.31:10.41&gt;</td>
<td>constraint priority</td>
</tr>
</tbody>
</table>

---

Cost=10:1.1:1.21:10.31:10.41>
Through the instructor interface, we feed the list of constraints specified by the teacher. The constraints are written into a DLV program, modeled as a constraint satisfaction problem as illustrated in figure 4. Here, we use two types of constraints: strong constraints and weak constraints (Buccafurri, 1997). The former are used to specify the conditions that have to be satisfied by the system in all cases. An example of these constraints would be that each student can be a member of only one group.

The weak constraints are used to specify the conditions that are preferably satisfied, but can be violated if the system would not be able to find a solution otherwise. These constraints are given a priority level according to their importance in the group formation through the instructor interface. For example, in figure 4, the instructor considers having only one leader (shaper or coordinator) in each group to be more important than having an implementer in each group by assigning these constraints priority levels 3 and 1 respectively.

Depending on the data provided and the constraints, DLV outputs more than one solution to the problem (i.e. more than one grouping of the students). Each solution is called a model. The optimal model is hence the best grouping of students in relation to the given constraints and input data. The best model is calculated as an objective function that minimizes the number of violated constraints.

Unlike other computer supported instructor-based group formation tools (Redmond, 2001) (Tobar, 2007), our approach does not leave any student orphans. Based on the negotiation of the constraints satisfaction through optimization, all students are allocated to some group, even if some constraints are violated. The best model is computed and the confidence of the computation (formation) is returned to the instructor: For instance, if the instructor wants only one leader per group to avoid conflicts, and gave the constraint priority 2, but the number of leaders is larger than the number of students; then some of the groups will have more than one leader. Here, a constraint of that priority is violated. Hence, the confidence of how good is the group formation is decreased. Together with the model, the confidence is computed in terms of violated constraints, and then returned as an output solution.

DLV outputs the model as a list of predicates. Figure 4 illustrates an example output predicates of the form “member(name,role,gender,group)” showing the students’ family name, Belbin role, gender, and the group they are allocated to. This output data is then stored in an SQL database and then returned to the instructor through the instructor interface as a list of groups. To ensure a good practice interface design, the interface output can be dynamically manipulated by the instructor in case modifications or swapping students around the groups is preferred.

**Evaluation**

As mentioned in the limitations of current systems, a well-defined approach to evaluating group formation has not been introduced yet. Hence, to evaluate the quality of the generated group formation, we defined a set of metrics that measure how good the formation is in terms of the constraints satisfaction rate of all the groups of students (Ounnas, 2007c). So far, we used the group formation framework with a range of strong and weak constraints and a different number of variables (i.e. students’ characteristics in which the formation is constrained). We used the framework to allocate students to groups within two courses in the University of Southampton including the software engineering course on the following year. However, since the instructors of these courses had only a maximum of three constraints (on previous marks, gender, and international students), the framework returned a best model in both cases with no violations in both courses.

To monitor the performance of the proposed framework, more evaluation scenarios were set. We run different scenarios of forming groups on simulated classes of 100 students. The simulated data was based on the population statistics collected from our observational study and confirmed from the UK Higher Education Statistics Agency (HESA: http://www.hesa.ac.uk/). The framework was tested with various constraints, different in content, number, and priorities. From the evaluation, we observed that our system can take up to 11 constraints before the solver starts taking a long time to process the groups. Although the results are still more efficient than any of the existing tools, the maximum number of constraints can be enhanced with the addition of some heuristics to the solver’s algorithm. With the heuristics implemented, the solver will be prevented from running out of time during the computation.
The group formation quality here is measured against the satisfaction of the constraints chosen by the teacher. This includes calculating the average of how many constraints have been satisfied for each group. The formation quality is then calculated in terms of and the average and standard deviation of the constraints’ satisfaction of each group, and therefore for the cohort in general. Formulas for calculating the formation quality are detailed in our evaluation framework, as described in (Ounnas, 2007c). In this context, we don’t measure the quality of the constraints themselves if they will lead to a good formation or not, neither do we take the students’ satisfaction with the allocations. Since teachers are given the freedom to choose the constraints and their importance, we trust they will choose the constraints that best fit their students’ needs and the collaborative task they are trying to achieve through the group formation. It is a part of our future work however, to evaluate the quality of different constraints in relation to different sets of data. Through many runs of experiments, the quality of the constraints can be measured as described in our evaluation framework (Ounnas, 2007c).

We also tested the system’s performance with incomplete data, by deleting data at random with an equal distribution on each characteristic considered for that group formation. Results showed that the systems still performs well (for example from no violations to 2 violations) when the data is down to 50% incomplete for a formation with 3 constraints, a moderate number of constraints for creating learning groups. However, as the number of constraints grows, the performance decreases accordingly when the data is incomplete. Table 3 shows the results of forming groups with incomplete data with a different number of constraints. For each scenario (with a different number of constraints), we calculated the number of constraints violations (NCV), and the group formation quality (FQ) as explained in our group formation evaluation metrics framework in Ounnas (2007c). The cells with no results show the case where the solver runs out of time during the computation.

<table>
<thead>
<tr>
<th>Data Completeness</th>
<th>With 2 constraints</th>
<th>With 3 constraints</th>
<th>With 7 constraints</th>
<th>With 9 constraints</th>
<th>With 10 constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCV</td>
<td>FQ</td>
<td>NCV</td>
<td>FQ</td>
<td>NCV</td>
</tr>
<tr>
<td>100%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>90%</td>
<td>2</td>
<td>0.91</td>
<td>2</td>
<td>0.94</td>
<td>3</td>
</tr>
<tr>
<td>70%</td>
<td>2</td>
<td>0.91</td>
<td>2</td>
<td>0.94</td>
<td>3</td>
</tr>
<tr>
<td>60%</td>
<td>2</td>
<td>0.91</td>
<td>2</td>
<td>0.94</td>
<td>3</td>
</tr>
<tr>
<td>50%</td>
<td>2</td>
<td>0.91</td>
<td>2</td>
<td>0.94</td>
<td>10</td>
</tr>
<tr>
<td>30%</td>
<td>3</td>
<td>0.86</td>
<td>6</td>
<td>0.74</td>
<td>11</td>
</tr>
</tbody>
</table>

These results were collected without using the semantic inference feature of the framework as it is still being implemented. We expect the results to improve, as the main use of the semantic inference is to handle incomplete data through the use of domain ontologies and deduction rules.

**Future work**

For future work of this research, we intend to test the framework with more challenging scenarios by increasing the number of students in the class, and the number and context of the constraints. Since groups can also be generated from social networks, a range of the constraints will be based on the social connections between the learners (established through FOAF relationships), such as forming groups of students who do not know each other directly, or forming groups of student who collaborated with each other in a previous task.

So far we only used the framework to run using complete data about the students. A more challenging task will be to generate groups from incomplete data. Since our framework is based on Semantic Web technologies, we intend to empower it to handle incomplete data using these technologies. For our future work, we plan to add a module to the architecture of the framework that mines data from web pages and connect it to the ontology and a set of deduction rules to infer the missing data from the knowledge base. In this case, if the student does not provide the data needed for the group formation constraints, then the system will substitute the necessary data and subsequently feed it to the DLV solver. For example, if the information about whether student John is a leader or not is missing, and we know from John’s web page that he is a captain of the football team, then we can infer that John is a leader; or if we are
grouping student by skills, and we don’t know Sarah’s skills, but we know that Sarah has a high grade in discrete mathematics and Sarah has a high grade in Logic then we can infer that Sarah will perform well in formal methods.

Once the framework is refined with deduction rules, the evaluation of its performance with incomplete data will be compared to its performance with complete data (and no deduction rules), and its performance with incomplete data (and no deduction rules).

Conclusion

In this paper, we discussed the challenges associated with the manual and automated formation of groups in Education by analyzing existing applications in the field, and running an observational study. Based on our findings, we proposed a framework for group formation based on Semantic Web technologies and constraint satisfaction optimization to assist teachers with effectively defining groups of students based on a set of constraints of their choice. Unlike existing Computer Supported Group Formation tools, the approach we followed does not leave any student unassigned to groups. Instead, we employ strong and weak constraints to negotiate the students’ allocations to groups and report the confidence of the generated solution. We evaluate the group formation based on the quality of the generated groups in terms of constraint satisfaction, and the robustness of the formation in case of incomplete data. For future work, we intend to evaluate the potential of employing Semantic Web technologies in compensating incomplete data through inference in deduction rules and domain ontologies.

References


Wilson, S., & Jones, P. R. (2002). *What is...IMS-IMS Learner Information Packaging? standards briefings series*. c. JISC.

InstanceCollage: A Tool for the Particularization of Collaborative IMS-LD Scripts

Eloy D. Villasclaras-Fernández¹, Julio A. Hernández-Gonzalo¹, Davinia Hernández-Leo², Juan I. Asensio-Pérez¹, Yannis Dimitriadis¹, Alejandra Martínez-Monés¹

¹GSIC-EMIC Group, University of Valladolid, Spain // evilfer@ulises.tel.uva.es // jhergon@ulises.tel.uva.es // juase@tel.uva.es // yannis@tel.uva.es // amartine@infor.uva.es
²University Pompeu Fabra, Spain // davinia.hernandez@upf.edu

ABSTRACT

Current research work in e-learning and more specifically in the field of CSCL (Computer Supported Collaborative Learning) deals with design of collaborative activities, according to computer-interpretable specifications, such as IMS Learning Design, and their posterior enactment using LMSs (Learning Management Systems). A script that describes such collaborative activities is typically designed beforehand in order to structure collaboration, and defines the features that determine the behavior of the LMS, for instance, the sequence of activities or the groups/role distribution. In CSCL settings, group management and composition are especially relevant and affect the chances of achieving the expected learning outcomes. This paper presents a software tool, named InstanceCollage, which aims at facilitating the configuration and population of groups for IMS-LD scripts created with the authoring tool Collage, and discusses the implications of the IMS-LD specification with respect to this task. InstanceCollage is designed to process collaboration scripts based on CLFPs (Collaborative Learning Flow Patterns). Using this type of patterns, InstanceCollage focuses on the importance of understanding the function of groups within the learning strategy of the script. This paper describes the approach taken in InstanceCollage to facilitate this understanding for non-expert users. Additionally, two case studies are presented, which represent complex authentic collaborative learning scenarios, as a proof of concept of the functionality of this tool. The case studies are also used to illustrate the requirements of group configuration tools and to show that InstanceCollage complies to such requirements.

Keywords

CSCL script, IMS Learning Design, group formation, Collaborative Learning Flow Patterns

Introduction

Computer Supported Collaborative Learning (CSCL) research aims at enabling new forms of interactions among students to foster learning. Computer support is central in distance collaborative learning environments, but it may be used also to scaffold face to face collaborative activities (Stahl, Koschmann & Suthers, 2006). Technology can play different roles in collaborative learning, including the enabling of interactions in distance settings, scaffolding or guiding collaborative interactions among participants, or providing resources to carry out learning activities (Suthers, 2005). This paper will look into computer support for the guidance of collaborative activities through scripts. In spite of the potential benefits, collaborative learning scenarios do not guarantee learning, as the chances of effective interactions among participants depend on a wide range of factors, including group composition, the environment and the characteristics of collaborative activities (Dillenbourg, 2002). With the objective of structuring and scaffolding the collaborative process, collaboration scripts have been proposed as an effective solution to create conditions in order to promote the chances of effective interactions.

Different types of scripts tackle the specification of different features of collaborative scenarios. Typically, collaboration scripts are classified as macro-scripts or micro-scripts (Dillenbourg and Hong, 2008). Micro-scripts describe the communication model between participants (Kobbe et al., 2007), such as argument construction processes. In CSCL settings, software tools that control the communication process and induce students to follow a micro-script have shown to enhance learning (Weinberger, Fischer & Stegmann, 2005). On the other hand, macro-scripts model learning objectives, the sequence of activities that participants have to follow, role distribution (Kollar, Fischer & Hesse, 2006), resource distribution, or group formation (Kobbe et al., 2007). Through the configuration of these components, macro-scripts define the pedagogical method of the collaborative activities that students are expected to perform. In a similar fashion as micro-scripts, macro-scripts can also be deployed in software systems, or LMSs.
The development of LMSs capable of automatically managing collaborative activities described in a macro-script is an active research topic in the CSCL community. This kind of LMSs can, according to the collaborative learning strategy that drives a macro-script, manage groups of students, control activity schedules and task distribution, and provide students with resources and collaborative software tools.

In order to adapt the LMSs behavior to the content of a script, formal specifications have been proposed. A formal specification for learning scripts provides a model and a computational representation to describe the features of a learning setting, which can be interpreted by software systems. In the case of macro-scripts, IMS Learning Design (IMS-LD) (IMS Global Learning Consortium, 2003), the de facto standard to model learning scripts, enables the description of activity sequences, groups/roles, task distribution, etc. Currently different authoring tools and players have been developed, facilitating the task of creating IMS-LD scripts and their enactment in software systems in real learning scenarios. Designed as pedagogically neutral, IMS-LD has also been used for model collaborative activities (Tattersall, 2006). In spite of certain limitations (Miao, Hoeksema, Hoppe & Harrer, 2005), such as the lack of mechanisms to describe complex activity sequences (Harrer & Malzalm, 2006), it has been argued that MS-LD provides sufficient support for relevant features of CSCL macro-scripts (Hernández-Leo, Burgos, Tattersall & Kopper, 2007; Koper & Burgos, 2005), and has actually been used as the computational representation format of CSCL scripts by the Collage authoring tool (Hernández-Leo et al., 2006a), as discussed in the following section.

This paper is focused on the configuration of a particular aspect of the enactment of CSCL macro-scripts in IMS-LD players: group particularization. We use the term particularization throughout this paper to refer to the configuration of groups (i.e., the specification of the necessary number of groups for each type of group, to which students are assigned) and the population of groups (i.e., the assignment of students to groups) for a specific scenario. Group particularization may occur in different moments, including during the creation of the script (the design phase), during the preparation of the script for a particular scenario (in a concrete LMS) or even during the enactment of the activities. We will discuss the implications of collaborative learning and IMS-LD with respect to this issue. Group management is especially relevant in collaborative learning (Dillenbourg, 2002), as it affects the development of activities, and thus the learning outcomes. Several research works (Ikeda, Go & Mizoguchi, 1997; Ounnas, Davis & Millard, 2008) have focused on the importance of creating optimum groups, in the sense that the characteristics of the selected students create better chances for each group's collaboration to result in learning.

However, group formation policies are also part of the mechanisms on CSCL scripts (Kobbe, 2006). Along a single script, students might be required to carry out activities in various groups: in different phases, they may work together with the whole classroom, in small groups, and even carry out individual activities embedded in the collaboration script (Dillenbourg, 2004). For instance, the ArgueGraph and Universanté scripts (Dillenbourg, 2002) or those based on the well-known jigsaw technique (Aronson, Blaney, Stephan, Sikes & Snapp, 1978), propose activity flows in which students must change groups, in order to carry out different activities. These group changes are sometimes key mechanisms of the rationale of the script: In the case of the jigsaw, students move from expert groups, whose members work on a specific topic, into jigsaw groups, composed by students that have worked on different topics; this mechanism is expected to trigger certain interactions. Therefore, group formation needs to be understood as part of the learning strategy on which the script is based, and is interrelated with activity sequencing and task distribution.

Accordingly, group formation should be reflected in the way the LMS guides students along the activities included in the script. Group formation, being part of the learning strategy of the script, is a requirement justified also in technological terms. For instance, management of document flow and of software collaborative tools requires knowledge about group configuration (Palomino-Ramírez, Martínez-Monés, Bote-Lorenzo, Asensio-Pérez, & Dimitriadiis, 2007). Computer support for the enactment of the script includes providing appropriate tools to enable the realization of activities such as communication tools (chats, forums, etc.) or others (e.g., collaborative editors) (Bote et al., 2008; Tchounikine, 2008). In this case, the system needs to take into account group configuration to allocate the needed resources and set appropriate access rights. For instance, the system should separate each group's private application space (for instance, chat sessions, own documents, etc.). On the other hand, document flow also depends on group composition. This is the case of, for instance, peer review activities, in which authors and reviewers need to exchange several documents in a precise way. If an LMS is expected to perform these tasks, the system needs information about groups: number, size and population.
These reasons have motivated the development of a software tool, *InstanceCollage*, with the objective of facilitating the particularization of groups that will intervene in a CSCL scenario. *InstanceCollage* is based on the general principles of a group formation tool, but it focuses on interoperability with IMS-LD compliant LMSs. While several tools currently allow users to particularize groups, they lack functionality to help non-expert users in the case of complex collaborative learning scripts, as we discuss in this paper.

The rest of the paper is structured as follows. The following section describes the type of script that *InstanceCollage* is focused on. After that, the paper discusses the task of group particularization (configuration and population, first in relation to the life-cycle of CSCL scripts in general, and later in the case of IMS-LD collaboration scripts. Having reviewed different approaches to this task, the paper describes *InstanceCollage*, indicating the rationale of its design as well as its main functionality. This functionality is further discussed in the context of three case studies, which illustrate the type of scripts that *InstanceCollage* is intended to process, as well as the technological issues related to group formation in authentic CSCL scenarios.

**Pattern-based approach for designing CSCL scripts**

*InstanceCollage* is proposed within the context of a particular approach to CSCL script design. Focusing on the difficulties that non-expert designers face when designing collaboration scripts, this approach is based on the application of a certain type of learning design patterns, Collaborative Learning Flow Patterns (CLFPs), to facilitate the design task (Hernández-Leo et al., 2006a). This section discusses this approach, which is a key element to understand the features of *InstanceCollage*.

Design patterns (Alexander et al., 1997), which have been applied to different fields, including architecture and computer science, have been also proposed in the domain of e-Learning, as a mechanism for sharing ideas about educational design problems (Goodyear et al., 2004) and validated general solutions. Among different types of design patterns for learning, CLFPs propose solutions to create CSCL macro-scripts in order to achieve certain learning outcomes, extracted from best practices in collaborative learning. Thus, CLFPs capture best practices in organizing collaborative learning sessions, which can be used by designers to identify possible solutions applicable to real scenarios. With respect to this, CLFPs can be applied to a wide range of scenarios, since they are neutral with respect to specific domain content. Examples of CLFPs are PYRAMID, THINK PAIR SHARE or the aforementioned JIGSAW (Hernández-Leo et al., 2006a).

In order to facilitate the design of CSCL macro-scripts for non-expert designers, an IMS-LD authoring tool has been proposed and used: *Collage* (Hernández-Leo et al., 2006a). *Collage* supports designers through the process of reusing and adapting CLFPs. This process involves the selection of one or more CLFPs, according to certain learning objectives, and the configuration of the embedded activities and resources to the characteristics of the scenario.

---

*Figure 1. Representation of a CLFP-based script*

*Figure 2. Groups and student change between groups in a JIGSAW-based script*
Additionally, this tool creates a computer-interpretable representation of the script, using the IMS-LD specification (Hernández-Leo at al., 2007a), thanks to which scripts generated with Collage can be interpreted with any IMS-LD compliant system.

As mentioned before, each CLFP captures a concrete learning strategy: each pattern aims at creating certain conditions to promote the occurrence of interactions among participants, which in turn are expected to promote certain learning outcomes. This is related to one of the main features of Collage: the usage of a graphical representation to depict the sequence and distribution of collaborative activities (also called learning flow). Each CLFP has a specific graphical representation, which uses graphical metaphors to facilitate the recognition of CLFPs, and to provide visual hints about the type of activities contained in the script. Figure 1 shows the graphical representation for CLFPs JIGSAW, THINK PAIR SHARE and PYRAMID.

A multicase evaluation study of Collage has shown that the use of CLFPs facilitates the design task, and that the tool itself provides an intuitive and easy way for non-expert users to design complex collaborative learning scenarios (Hernández-Leo et al., 2008). However, as discussed in the following section, the IMS-LD specification, and therefore Collage, does not provide means to allow assigning students to groups. In addition, scripts created with Collage may have a complex group structure, rooted in the activity sequence and involving several group changes. Thus, configuring and populating groups in a real scenario can be a challenging task. Figure 2 shows how students change groups between the three phases that compose a JIGSAW-based script. The complexity of the script increases the odds of making mistakes in group formation, which would consequently hinder the development of the activities. For instance, a typical error may occur if a jigsaw group does not contain an expert from every different topic, as JIGSAW requires. This type of requirements needs to be taken into account during group particularization.

**Approaches to the configuration of groups in CSCL settings**

In the previous section we indicated that InstanceCollage is focused on scripts generated with Collage, that is, IMS-LD scripts based on CLFPs. This section discusses the technical implications of IMS-LD with respect to group particularization. Before focusing on the characteristics of IMS-LD, we will review the life-cycle of CSCL scripts. Finally, this section reviews the particular characteristics of CLFP-based IMS-LD scripts.

**Life-cycle of CSCL scripts**

From the conception of a CSCL macro-script to its application in a real scenario, there exist different phases of a CSCL script life-cycle that involve several actors. Even though there is no consensus about the exact composition of the life-cycle of CSCL scripts, it is possible to find several proposals in the literature. For instance, Weinberger, Collar, Dimitriadis, Mäkitalo-Siegl, & Fischer (2008) proposes a life-cycle, emphasizing the different steps involved in the production of an effective CSCL script, before its application in a computer-supported environment. This life-cycle includes the specification of the script, its formalization with a computer-interpretable specification and, possibly, a simulation of the activity flow, before deploying the script, i.e., enacting the activities with students. On the other hand, other research works have focused on the technical issues around the enactment of CSCL script in LMSs. Vignollet, Ferraris, Martel, & Burgos (2008) describe a life-cycle with the following four phases: design, operationalization, enactment and evaluation. Hernández-Leo et al. (2006b) describe a similar approach, shown in Figure 3, using the term instantiation (IMS Global Learning Consortium, 2003) to refer to the phase in which the CSCL script is prepared to be enacted. In both cases, the operationalization and instantiation phases include the tasks necessary to prepare the script for enactment in a concrete scenario using an LMS. The rest of this paper will employ the term instantiation in the sense of creating an instance of a CSCL script for a specific situation.

We are interested in the relationship between design, instantiation and enactment, with respect to group particularization. Configuring and populating groups can be accomplished during the design phase. In this case, the designer may use information about the participants to create the script. For instance, Isotani & Mizoguchi (2007) propose a tool, Chocolato, which considers not only the number of participants but also their current learning state, in order to provide suggestions concerning the most appropriate configuration of collaborative learning activities. Information about participants can be, therefore, valuable for designers (and tools with intelligent support for design) to select the pedagogical method on which the script is based.
However, scripts do not contain necessarily all the information needed for the LMS to manage the activities. In particular, group composition may not be defined in the script. This approach is taken by the IMS-LD specification, which does not allow the definition of group population, as discussed below. Similarly, the learning design authoring tool developed by the LAMS project (Dalziel, 2003), does not include the task of populating groups during the design phase. In the case of LAMS, group population is done during enactment. The designer can however specify the group population policy (random, student choice or staff decision), which will determine the behavior of the player. Excluding group population from the script definition makes the script independent of a particular learning situation, and thus facilitates its reutilization. However, it also demands further configuration in later stages after design. Next section discusses this issue for the case of IMS-LD collaboration scripts.

**Current approaches in IMS-LD systems**

If we focus on the case of IMS-LD scripts, it is the specification itself which constraints the moment when the particularization of groups is performed. As we indicated before, an IMS-LD script does not indicate the students assigned to each group. Therefore, the population of groups must be done always after the design phase: in the instantiation phase, or even during enactment.

With respect to groups, IMS-LD provides support to their definition by means of the role concept (IMS Global Learning Consortium, 2003). This concept is used to assign activities to different students: the script specifies the activities that each role performs and, in a particular situation, the participants are assigned to different roles. When participants are assigned to the same role, then they form a group. In collaborative learning, we are interested in modeling the existence of different groups, among which students are distributed. In order to model these groups, we need two “levels” of group information:

a) In a single script, different types of groups may exist. Usually, different types of groups intervene in different phases, and each type can be used to distribute students in groups in a different manner. In the case of IMS-LD, different types of groups can be specified in the script; therefore, they are created in the design phase. For example, in the JIGSAW shown in Figure 2, there are two types of groups: jigsaw and expert groups.

b) For each type of group, several groups may exist. In this case, these groups of the same type represent groups of students working in parallel in the same activities. Figure 2 shows four instances of jigsaw groups and three instances of expert groups.

With respect to having different groups of the same type, IMS-LD allows two approaches:

1) For each group of students, a different role is created at the design phase. In this case, the designer needs to know the exact number of groups that will exist in the classroom. When different groups perform the same activity in parallel, IMS-LD allows assigning those groups to the same activity. Using this approach, during the instantiation phase it is only necessary to assign students to groups.

2) For each different type of group, only one role is designed. In this approach, several roles may be created to represent the different types of groups that students will take during the script. However, in order to have the necessary number of groups of each type (to allocate students in different groups), they have to be created in the instantiation phase. Therefore, in this approach the design in simpler (the exact number of groups does not need to be known in advance), but instantiation requires also the configuration of the number of groups.
While both approaches are possible according to the IMS-LD specification, it seems clear that the later is better aligned with the option of leaving the population of groups for the instantiation phase. In contrast to the first approach, the second facilitates the reutilization of the script in scenarios with different number of participants, since the number of groups is not predefined in the script.

**Issues concerning CLFP-based scripts**

Continuing with the discussion of the previous section about IMS-LD, now we focus on the specific case of CLFP-based scripts (i.e., those created with the Collage tool).

Previously we indicated two possible approaches to create different groups (of the same type) of student in IMS-LD scripts. Collage follows one approach or the other depending on each CLFP. Actually, only one CLFP allows the creation of different groups during the design phase: SIMULATION. Only in this case Collage offers functionality to create several groups, to which different names and activities can be assigned (including, therefore, different resources). The SIMULATION has this special feature because it was considered that the definition of the simulation learning process itself needs the identification of all the intervening roles (which can represent an individual or a group).

On the contrary, in the case of all the other CLFPs Collage takes the approach of defining one single role for each group type, thus leaving the creation of the necessary copies for the instantiation phase. For those CLFPs, therefore, Collage does not offer functionality to alter the number or groups. For instance, in the case of the JIGSAW, the script defines only one jigsaw group and one expert group. The PYRAMID, on the other hand, defines exactly one group per pyramid level. In order to distribute students among groups, it is necessary to create additional jigsaw, expert, etc. groups during the instantiation phase.

Creating the necessary groups can be a challenging task. While each CLFP defines (except for the SIMULATION) a low number of group types, scripts generated with Collage may contain several CLFPs. CLFPs are combined hierarchically, in the following way: one CLFP (we will call it parent CLFP) defines the global structure of the script. Other CLFPs (child CLFPs) can then be used to structure the activities of each phase of the parent CLFP. When this happens, the groups defined in the child CLFP become subgroups of each group that intervenes in the affected phase of the parent CLFP. With respect to the instantiation phase, this mechanism requires that the groups defined in the child CLFP be configured as many times as the number of groups of the parent CLFP. For instance, in Figure 1 the TPS and JIGSAW are each embedded in a phase of the PYRAMID. TPS groups, therefore, would need to be configured for every group of the level 1 of the PYRAMID.

Finally, it is important to note that each CLFP may have its own rules, derived from the learning strategy on which the pattern is based. Figure 2 shows the rule implicit in the JIGSAW CLFP: groups formed in one phase must split so that each member should join a different group in the following phase. Another example is the Thinking Aloud Pair Problem Solving (TAPPS) (Lochhead & Whimbey, 1987), which requires creating pairs and, for each pair, distributing solver and listener roles explicitly. Assigning these roles is necessary, since they determine the activities performed by each member of the pairs.

In order to take all the above issues into account, and especially in the case of complex scripts (for instance, when several CLFPs are used), the person responsible for the particularization of groups could benefit from information about the exact purpose of each group in the script, including the CLFP to which each group is related, as well as the activities assigned to the group. Next section explains how InstanceCollage provides users with this information in the same tool that is used to configure groups.

**Particularization of groups in InstanceCollage**

Based on the positive evaluation of Collage, the importance of group formation, and the relation between groups and the learning strategy of CLFPs, InstanceCollage was proposed to facilitate the task of group particularization. While Collage supports teachers in the creation of CLFP-based scripts, InstanceCollage is focused on the instantiation phase, aiming at providing an effective and efficient solution for this complex task. This solution is based on the fact
that CLFPs also contain information about group formation policies, and thus they can be used to provide valuable support for the instantiation phase.

*InstanceCollage* produces a description of the groups and their population, which is applicable only for the case of the specific instantiated CSCL script. The description is stored in XML format. Additionally, *InstanceCollage* can upload the particularization information to an LMS. Currently *InstanceCollage* is capable of publishing a CSCL script in *CopperCore*, creating a session and particularizing groups according to the configuration created by the user. In this way, the tool is integrated more transparently in the life-cycle of the script. Additionally, interoperability with another IMS-LD interpreter, named *GRAIL* (del Cid, de la Fuente-Valentin, Gutiérrez, Pardo, & Kloos, 2007), embedded in the open-source .LRN platform, is currently underway.

**Rationale**

*InstanceCollage* facilitates the particularization of groups by representing them in relation to their function in the script. To achieve this, *InstanceCollage* uses the same graphical metaphors employed by *Collage* to represent each CLFP. Figure 4 shows, for instance, the graphical representation of the THINK PAIR SHARE CLFP. In this way, the user may be familiarized with the metaphors of each CLFP. If this is not the case, still the graphical representation is expected to be easy to understand. The metaphors are intended to be simple and, rather than displaying complete information about the script, they provide a hint about the structure of the collaborative activities.

![Figure 4. Creation of groups in InstanceCollage](image)

The graphical representations for CLFPs depict, at least, phases and groups. Phases are typically represented vertically, separating the different steps of each CLFP. Additionally, the graphical representation uses specific symbols to represent the groups that intervene in each phase. In the case of the JIGSAW CLFP, for instance, colored jigsaw pieces represent different types of groups; in the SIMULATION CLFP, different icons represent each intervening role. In addition to using a meaningful representation for each group, the graphical interface locates
groups within the activity flow, indicating the activities in which each group intervenes. In this way, \textit{InstanceCollage} facilitates identifying each group's function in the script, and understanding the relation between each group and the rules for group formation derived from CLFPs. In this way, \textit{InstanceCollage} connects the learning strategy of the script with concrete information (i.e., the identity of participants) of the scenario in which that script will be enacted.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Population of groups in \textit{InstanceCollage}}
\end{figure}

\textbf{Usage}

In order to complete the task of group particularization, \textit{InstanceCollage} guides its users through four steps, that are presented in different tabs after loading in the tool an IMS-LD script created with \textit{Collage}. The first step, shown in Figure 4, consists in the creation of the required groups. To do this, the user can browse through the different CLFPs and their phases. After locating the group that intervenes in each phase, the user can create the number of groups necessary; when the configuration begins, each group has exactly one copy. Figure 4 represents the creation of two \textit{pair} groups (for a total of three), using the menu option “clone”. On the left side of the figure, a tree shows the CLFP structure and the intervening; three \textit{pairs} can be seen both in the graphical representation and the tree structure.

The second step (the first and second steps can actually be performed in any order), involves simply retrieving the names of the participants, including teachers and students. Currently \textit{InstanceCollage} loads this information from a text file provided by the user. Once \textit{InstanceCollage} has information about groups and students, the user can assign students to groups. The third step corresponds thus with the population of groups, and it is represented in Figure 5. The presentation is similar to that used in the first phase, but now additional information is shown concerning the users that belong to each group. In addition to intuitive means to add or drop users to the different groups, for each phase the tool indicates which users have been already assigned to a group, and those who have not, with the objective of preventing mistakes. In Figure 5, the graphical interface shows that \textit{Student_1} and \textit{Student_2} have already been assigned to a \textit{pair} (using a different color in the user list on the right-bottom); the user then adds the next two students to the second \textit{pair}. 

63
In order to facilitate the first and third steps, the graphical representation is interactive. Each type of group is represented with a different icon, and the number of icons represents the number of groups for each type. By interacting with the different elements of the representation, groups can be created or deleted (thus modifying the number of groups of the same type), and students can be assigned to these copies (thus populating groups).

The final step involves the deployment of the script in the LMS. InstanceCollage automatically uploads the script; then, groups are created and populated for the particular scenario, according to the group particularization created by the user. After this step has concluded, the script is ready to be played by participants in the LMS. The interoperability between InstanceCollage and the LMS can be bidirectional: first, InstanceCollage might retrieve the list of participants, thus simplifying the second step described here. Second, InstanceCollage needs to connect with the LMS to upload the information about groups.

Case Studies

This section discusses two case studies in which CSCL scripts created with Collage and formalized with IMS-LD were applied in learning scenarios. These case studies represent the whole life-cycle of real CSCL scripts in LMSs, including the particularization of groups depending on the participants (and their characteristics, such as their physical location) of authentic learning scenarios. At the moment in which each of these cases studies were carried out, different software tools were used to complete the group particularization. Afterwards, we carried out the particularization phase again using InstanceCollage. Our objective is, first, to evaluate whether InstanceCollage is capable of replicating the group particularization as was required in each case study. Additionally, this section examines the difficulties in the group particularization caused both by script complexity and the educational context. These difficulties are discussed with respect to the tools used in the case studies and the potential benefits of InstanceCollage.

Case Study A: CTM2

The first case deals with a fifth year undergraduate course of Telecommunication Engineering, at the University of Valladolid, with 12 students attending the course. The case study comprises two two-hour sessions in which the students have to read a complex article on the topic of Network Management and analyze the main ideas presented in it (Hernández-Leo et al., 2007b).

In order to foster discussion among the students about the paper, a CSCL script was planned. This script aims at guiding students along a series of activities that involve individual learning, discussion, giving explanations to peers, or reaching agreement. The CSCL script was created with Collage, based on three patterns: JIGSAW, PYRAMID and TPS, organized as shown in Table 1. On the other hand, the Gridcole system (Bote et al., 2008) was employed for the CSCL script enactment, using a player based on the CopperCore IMS-LD engine. Gridcole implements a service oriented architecture, which enables the tailorability of the technological support for the learners by the integration of third-party tools (collaborative or not) on-the-fly. Other IMS-LD players, such as SLeD (Weller, 2006) could have been used. SLeD actually provides a graphical interface for the group particularization, more user-friendly than the tool provided in CopperCore. However, the case study was also used to evaluate the integration in Gridcole of third-party services; this was the main reason for the staff to choose this player. Actually, several tools, offered as services, were used by the participants in the case study: Synergeia, a shared, structured, web-based work space for collaborative learning, which includes tools for document sharing and asynchronous discussion; a chat, for synchronous communication; and Quest, a tool to publish and answer questionnaires on-line. As discussed below, there is a strong relationship between service allocation and provision and group particularization. This issue is currently being researched in Gridcole, which supports the binding of different instances of these services-like tools to each group in each activity. In this way, each group is provided with its own private work space for each tool. This binding is, therefore, based on the group particularization. For this reason, it was necessary to configure all groups before enacting the activities.

At the time when the case study took place, groups were particularized using Clicc, the tool provided with CopperCore. The particularization required 56 actions; being console-based, each action in Clicc is performed by means of a textual command. In spite of the low number of students, a total of 13 groups in total had to be created,
and each student had to be assigned to 4 groups (of different type), besides the class group that included all students. Figure 6 represents the group structure created for this experience. The usage of Clicc presented important difficulties. In addition to requiring the user to learn a series of text commands, the information provided by the tool concerning the particularization of groups as it was completed was difficult to read: it is presented in a XML-like text description, heavily based on the IMS-LD specification, which might not be readable for non-expert users. Actually, the configuration of groups was decided before using the tool with “pen and paper”, in order to understand the 56 commands necessary.

The same group particularization, with the same students, was carried out afterwards with InstanceCollage (Figure 7 shows the creation of the necessary groups). Loading the resulting particularization created with InstanceCollage in Gridcole yields the same result as the particularization with Clicc. However, the benefit of InstanceCollage lies in the guidance provided by the graphical interface. This tool provides comprehensive information to the user, which can be used to analyze and decide the number of groups and their population. In contrast to Clicc, InstanceCollage does not only enable the configuration of groups, but aims at supporting the user in the process of taking the decisions that complete this task.

Case Study B: Mosaic

The second case was carried out in a coordinated activity between three universities, in the context of graduate courses on the topic of grid computing. Twelve students from the University of Valladolid, the Open University of Catalonia and Carlos III University of Madrid (four from each institution), participating on independent doctoral courses, worked together following a collaborative script. Given that each student worked from their own institution premises, computer support was particularly important to carry out the collaborative activities, which included both synchronous and asynchronous ones. The CSCL script was enacted using GRAIL (del Cid et al., 2007). One teacher from each institution monitored the activities locally.

The script was, similarly to that of Case Study A, composed by the PYRAMID and JIGSAW CLFPS, and PEER REVIEW activities; group particularization was done almost in the same way, but without configuring TPS groups. In this case, PEER REVIEW activities were included in different phases: first individually, during the JIGSAW (in the first level of the PYRAMID), and later in groups, during the second level of the PYRAMID.

The pedagogical context highlights the importance of providing information about the learning strategy of the script to facilitate group population. One of the distinguishing characteristics of this case study is the involvement of three teachers in the activities. After one of the teachers created the script with Collage, a second teacher configured the LMS to support the document flow between participants, which was done semi-automatically: the tool provided personalized instructions to each student about how to retrieve and upload the documents. Configuring the support for document flow requires information about groups and the activity flow. In different phases, the roles of authors and reviewers were distributed in different manners: for instance, exchange of documents happened between individual students or whole groups, depending on the phase. Therefore, document flow had to be configured taking into account not only the groups, but also the related activities. Moreover, the population of groups was decided
considering the physical location of each student, in order to foster students to collaborate both with students of the two other institutions and their own.

Figure 8. GRAIL interface for the population of groups

In the case study, group particularization was performed with the tool provided by GRAIL. Figure 8 shows a screenshot of this tool, in which the graphical interface can be seen. By using a web interface, the tasks of creating and populating groups are more intuitive than in Clicc. However, some problems were detected:

- When configuring different group “copies” (for instance, the three expert groups shown in Figure 8), the tool does not offer information about which participants have already been assigned to one of the groups of the same type. This can increase the chances of making the mistake of assigning the same student to several groups of the same type.
- In the case of group hierarchies (i.e., when parent groups contain child groups), the tool does not allow the assignment of students to the parent groups automatically by assigning them directly to child groups. In addition, there are no visual indications about the complete group structure.

Both issues are tackled in InstanceCollage. The particularization of the script used in this case study in InstanceCollage shows that it is possible to replicate the same group structure and population. In addition, InstanceCollage indicates, for each type of groups, what students have already been assigned to one group, in order to prevent the error of assigning one student to, for instance, several expert groups.

Discussion and comparison with other tools

In the two cases, the collaborative activities were predefined in a CSCL script, based on CLFPs. An LMS was used to support the enactment of the activities. Complementing Figure 6, Table 1 shows the CLFP structure that defined the sequencing of activities. In the original cases, Clicc and GRAIL were used to particularize groups.

For both cases, InstanceCollage has shown to be capable of replicating the result of the group configuration and population phase. Since the number of students is low, further tests with the tool were carried out, showing that InstanceCollage supports the particularization of scenarios with a higher number of students. Up to 100 students were assigned to groups during the particularization of a test script; the group structure and population was later successfully uploaded in CopperCore.
Table 1: Necessity of support for group configuration and population.

<table>
<thead>
<tr>
<th>Case study</th>
<th>Script</th>
<th>Need to show the learning strategy</th>
<th>Technological issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) CTM2</td>
<td>PYRAMID</td>
<td>Complexity of the script (group hierarchies).</td>
<td>Several tools provided by the LMS. Group resource allocation depends on the group configuration.</td>
</tr>
<tr>
<td></td>
<td>Level 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JIGSAW</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Level 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Level 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TPS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B) Mosaic</td>
<td>PYRAMID</td>
<td>Complexity of the script (group hierarchies).</td>
<td>Document flow depends on group formation:</td>
</tr>
<tr>
<td></td>
<td>Level 1</td>
<td></td>
<td>- Types of groups (individual or group roles)</td>
</tr>
<tr>
<td></td>
<td>JIGSAW</td>
<td></td>
<td>- Specific participants.</td>
</tr>
<tr>
<td></td>
<td>Ind. PEER REVIEW</td>
<td>Different actors might be involved in the group configuration task.</td>
<td>- Activities in which document flow takes place.</td>
</tr>
<tr>
<td></td>
<td>Level 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group PEER REVIEW</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Level 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The tools originally used at the time of the case studies, not being specific of collaborative learning, show problems when configuring scripts in which students form complex group structures. In comparison with other tools, *InstanceCollage* is the only one that provides visual information about the moment in which groups intervene and, therefore, their function in the learning strategy of the script.

*Clicc* offers an unintuitive interface, which requires the user to enter a series of text commands. The information that this tool provides to describe the group particularization as it is been configured is presented in XML format, which makes its understanding difficult. On the other hand, the application embedded in *GRAIL* offers a more user-friendly interface. This tool implements easy-to-use group creation and population functionality. However, in the case of CSCL scripts in which both the activity sequence and group structure are relatively complex, the tool lacks the functionality of automatic assignment of participants in child groups to the parent group and of providing information about student distribution among groups during the population task.

Finally, there are other available tools for the instantiation of IMS-LD scripts, typically integrated with IMS-LD interpreters. *SLeD*, another player based on the *CopperCore* IMS-LD engine, also provides a graphical interface to particularize groups. In contrast with *GRAIL*, *SLeD* shows in a tree representation of the group hierarchical structure, facilitating the overview of the groups (types and copies of groups). However, similarly to *GRAIL*, it does not provide information about the participants already assigned to groups of a certain type. The *ReCourse* IMS-LD editor includes a *CopperCore* manager that enables the population of roles (i.e., groups). However, this tool does not allow the creation of different copies for each group type during instantiation. Therefore, this tool cannot instantiate scripts generated with *Collage*.

In contrast to other tools, *InstanceCollage* acknowledges the complexity of collaborative learning settings, such as the difficulty and importance of distributing students in groups adequately, and in accordance to the learning strategy, especially when groups are used as a means to improve the chances of social interactions. Information from CLFPs is used to display and communicate information about the purpose of groups within the script, not only during design, but also in the instantiation phase.

However, *InstanceCollage* presents a series of limitations. The first and main limitation is derived from the usage of CLFPs, which determine that only CLFP-based IMS-LD scripts can be processed by this tool. This limitation is shared with the related authoring tool, *Collage*. Therefore, *InstanceCollage* is not intended to substitute completely other group particularization tools, which are still needed to instantiate scripts generated with other authoring tools. However, we argue that *InstanceCollage* provides a more intuitive graphical interface and functionality in the case of CLFP-based scripts. CLFPs have prove useful to facilitate the access to collaborative learning scripts for non-expert users (for instance, teachers), and therefore *InstanceCollage* can help such users to close the life-cycle of CSCL scripts.
Moreover, a second limitation appears as the number of student increases. While InstanceCollage can handle a relatively large number of students, group population may become cumbersome, since this task needs to be done for each student (each of them may be assigned to several groups). Actually, this problem is shared by all the role/group particularization tools known to the authors. InstanceCollage, however, opens the possibility of implementing algorithms for automated group population that are aware of the pedagogical method of the script. Actually, there exist in the literature several approaches to automated grouping of students, based on different grouping criteria, for instance those proposed by Ikeda et al. (1997) and Ounnas et al. (2008). In addition to such strategies, CLFPs provide information about how to configure groups; see for instance the case of the JIGSAW in Figure 2. This information can be used to implement algorithms that automatically create groups that respect the pedagogical method of the script, depending of the CLFPs employed.

A third limitation is a technical one: in order to deploy the script, InstanceCollage needs to interact with the IMS-LD player. Currently InstanceCollage can automatically upload script (including the creation and population of groups) only in CopperCore and Gridcole. As mentioned before, the interoperability between InstanceCollage and other players, such as .LRN, is under development.

Conclusions

This paper has described InstanceCollage, a tool for the configuration and population of groups in CSCL settings. Focused on IMS-LD scripts, InstanceCollage adopts a pattern-based approach, in which CLFPs are used to provide users (instructional designers or collaborative learning practitioners) with information about the pedagogical method of CSCL macro-scripts. InstanceCollage is intended to process scripts generated with the authoring tool Collage. In spite of the limitations due to the usage of CLFPs, these patterns and their combination offer important advantages to support the particularization of groups for non-expert users.

IMS-LD has shown to provide adequate support for collaborative learning scenarios. To that end, the life-cycle of scripts formalized with this specification requires further configuration of groups after the design phase, as this configuration will affect the development of activities. The two case studies discussed in this paper show that group formation is critical in CSCL settings. Besides guiding the enactment of activities, computer support includes the provision of software tools that enable collaboration, such as communication tools, and management of document flow. These technical issues justify the interest in creating detailed descriptions of group configuration and population, since implicit grouping may not be sufficient for the LMS to manage the delivery of meaningful collaborative activities.

Therefore, tools that enable users to particularize groups are needed. The case studies show that InstanceCollage is capable of configuring two CSCL macro-scripts which have been used in authentic situations. These scripts, based on CLFPs, are characterized by a complex pedagogical method. CSCL scripts may require assigning each student to several groups, using group changes between phases to create conditions that promote interactions. In addition, considering the technical issues and the whole life-cycle of CSCL scripts, it is possible that several actors (teachers, pedagogy experts, or other staff) are involved in the process of applying a script to a learning scenario. These different actors need to exchange information about, among other things, the learning strategy of the script and group formation decisions concerning the particular scenario. This issue reinforces the need to support the understanding of this information in the instantiation phase, for which InstanceCollage was developed.

In order to facilitate the particularization of groups according to the learning strategy of the script, the responsible staff or teacher should be aware of the relation between groups, activity design and the pedagogical method of the script, since grouping policies and related mechanisms (e.g., group changes) are key elements of the activity flow. Providing this information, in an intuitive graphical interface, is the main contribution of InstanceCollage with respect to other existing tools. This is achieved by means of CLFPs and their pattern-specific graphical representation. In spite of the limitations of InstanceCollage, the approach implemented in this tool for group particularization can be specially beneficial in the case of non-expert users working with complex collaborative scripts.

Further evaluation will be aimed at investigating the intuitiveness and ease of use of the tool. On the other hand, future work will be aimed at implementing automatic group population, which may be especially valuable in cases...
with a large number of students. Automatic group population, however, needs to take into account the pedagogical model of each CLFPs. JIGSAW and SIMULATION patterns feature complex group changes; in the PYRAMID, on the other hand, appropriate group size in each phase is a relevant factor. Rules derived from CLFPs can be combined with other group formation policies based on student characteristics. Finally, it is worth noting that current work deals with the integration of InstanceCollage with LMSs capable of interpreting IMS-LD scripts, focusing on the automation of collaborative tools management.

References


Toward a Semantic Forum for Active Collaborative Learning

Yanyan Li\textsuperscript{1}, Mingkai Dong\textsuperscript{2} and Ronghuai Huang\textsuperscript{1}

\textsuperscript{1}School of Educational Technology, Beijing Normal University, China // liyy1114@gmail.com // Huangrh@bnu.edu.cn
\textsuperscript{2}Corporate Technology, Siemens Limited China, Beijing, China // mingkai.dong@siemens.com

ABSTRACT

Online discussion forums provide open workspace allowing learners to share information, exchange ideas, address problems and discuss on specific themes. But the substantial impediment to its promotion as effective e-learning facility lies in the continuously increasing messages but with discrete and incoherent structure as well as the loosely-tied learners with response-freeness. To motivate and facilitate active collaborative learning, this paper describes the design of a semantic forum with semantic link networking on discussion transcripts. Based on domain ontology and text mining technologies, messages are automatically processed for structural modeling with semantic association and special interest groups are automatically discovered for topic-centric social context measurement, which lays the foundation for the fulfillment of distinctive functionalities in the semantic forum (i.e. semantic search, relational navigation and recommendation). Compared with traditional forums, the semantic forum has three outstanding features. First, it deals with the structural incoherence and content isolation within forums. Second, it enables active learning by providing learners with relational navigation to meet their learning demands. Third, it supports social context based ranking to recommend learning companions or transcripts for collaborative problem-solving.

Keywords
Semantic forum, Social context, Semantic link network, Relational navigation

Introduction

Proponents of collaborative learning claim that learners in cooperative teams achieve higher levels of performance and retain information longer than learners who work individually (Webb, 1995). Asynchronous discussion forum plays an important role in supporting collaborative learning, allowing learners to communicate at large to ask questions, articulate their thoughts, explain and justify their opinions, share ideas and resources, with collaborative contribution to the knowledge building within online discussion forum. The accumulated discussion transcripts contain a great deal of deposited human knowledge. So, almost all e-learning platforms or environment provide discussion forums, expecting to support flexible collaborative learning and maximize resource sharing and reuse.

However, there exist some challenges and difficulties that hinder online discussion forums being as an effective e-learning facility. Firstly, the messages being delivered are arbitrarily structured and the content is dynamically generated through multiple learners’ inputs, interactions, and annotations, thus learners have to spend a lot of time to go through nonlinear messages for locating useful messages. Second, messages are incoherent and disordered without semantic association. So, learners cannot search for the specific topic by retrieving certain discussion transcripts in separation. Last but not least, learners usually feel isolated and are easily tended to become disengaged and inactive when they are outside of the social context of the classroom (Amy et al., 2002). The literature in education research suggests that learners who are actively engaged in learning will be more likely to achieve success (Bonwell, 1996)(Richards, 1995)(Rubin et al., 1998). But investigators have come to recognize that asking learners to interact and discuss on forums does not necessarily lead to high-quality discourse. So, how to motivate learners to be actively engaged in collaborative learning via scaffolding intelligent facilities in forums becomes an important issue.

Therefore, by incorporating domain ontology and text mining technologies, this paper proposes a semantic forum that focuses on the semantic organization of discussion transcripts with semantic associations. By analyzing the content of messages as well as interaction between learners, an approach is also presented to discover special interest groups (SIGs) on the same or similar themes for topic-centric social context measurement. Meanwhile, a semantic search with relational navigation support is also provided to guide the learners through well-structured and coherent messages relevant to their learning demands. With such facilities, online discussion forum not only supplies communication space, but also serves as full-fledged learning environment with evolved structured knowledge repository to better support highly active and peer-to-peer in-depth collaborative learning.
Related Works

Analysis of the relationship among participants in discussion forums draws a lot of attention from researchers (Hewitt et al., 2007). Social Network Analysis (SNA) is an established method to derive person-person relations in the form of sociograms from "traces" of communication in a networked community (Wassermann et al., 1994). It is widely used to study the way people participated and interacted with each other in discussion boards (Laat et al., 2005), which provides information about the activities of such a community and the way they learn collaboratively. Based on writing and responding to messages, researchers have concentrated particularly on the constructed social networks and properties such as the small-world property and power-law degree distributions. This method is simply based on the information flow between learners but ignores the content of messages, and is therefore unable to discover interest groups to support theme-centered in-depth communication and learning. Wang et al. (2006) propose an analysis framework composed of forum level, thread level, and member level for assisting investigations on forum communities, in which the member level analysis is to segment members into different categories and identify their roles. Kato and Akahori (2005) study the familiarity among participants on a BBS by conducting an experiment with focus on self-disclosure of participants.

As a matter of fact, content of messages in the discussion boards are important for studying the relationship between learners. Researchers in the field of Computer Supported Collaborative Learning (CSCL) focus upon using content analysis technique to analyze transcripts of asynchronous discussion (Laat et al., 2005)(Schrire, 2006). Wever (2005) summarizes various content analysis schemes to analyze transcripts of online asynchronous discussion. Researchers usually annotate the messages one by one either by manual-coding or depending on assistance of content analysis software, which cost much time and efforts. So, this method only applies to small amount of messages.

Alternatively, some researchers put forward automatic method to discover useful information from online discussion forums (Dringus et al., 2005). Murakami et al present a method to extract information from web discussion boards by summarizing threads (Murakami et al., 2001). They use quote and comment relationships to extract a thread summary, which indicate there are topic bindings between messages. But as the messages on the forum are discrete and incoherent, so how to discover the topic-related messages and structure them orderly has to be solved. Xi et al. (2004) propose a specialized ranking function for Usenet by using linear regression and support vector machine techniques. Their approach is based on metadata, such as prior knowledge about the message author or the depth of the message. They do not address short message problem. Kim et al. (2005) propose a topic segmentation approach to discover the topic divergence within one thread. This knowledge can be used in segmenting the content in coherent units and guiding the learners through segments relevant to their navigational goals. But actually, the average thread depth is actually low in many discussion forums, so we emphasize the semantic organization of large amount of messages and support for semantic navigation instead of analyzing one discussion thread.

The Framework for Designing a Semantic Forum

The designed framework of a semantic forum illustrated in Figure 1 includes the following core modules:

- **User Interface**, which allows the geographically distributed learners to access the forum via a friendly interface. After successfully logging in, learners can asynchronously discuss with other learners via interaction facilities.
- **Text Corpus**, which saves all the original discussion transcripts delivered by the participating learners.
- **Knowledge Map (KM)**, which expresses the concept map of the discourse (i.e. concepts and the relations between concepts) by referring to domain ontology. Herein knowledge map can be drawn by experts or advanced learners via the authoring tool.
- **Transcripts Processing and Organization**, which is in charge of analyzing the discussion transcripts based on text mining technologies. Via such module, the transcripts are structured with semantic association.
- **Knowledge Engineer**: who is responsible for assisting transcripts processing and maintaining the knowledge repository.
- **Knowledge Repository**, which saves the processed messages with annotations of topic and type in a structured and semantic-associated manner. It dynamically evolves and serves as an objectification of the community’s advancing knowledge.
- **Intelligent Facilities**, which offers various functions to support active collaborative learning.
Figure 1. The framework for designing a semantic forum

**Processing Discussion Transcripts for Semantic Organization**

**Message Topic Recognition**

Assuming that the messages in a discussion thread represent the same topic, we combine them into a summary document and then process it. Each document is represented as a weighted term vector \( d = (d_1, d_2, \ldots) \) with the standard TFIDF function.

\[
d_i = TF(w_i, d) \cdot \log \left( \frac{|E|}{DF(w_i)} \right)
\]

(1)

Where the term frequency \( TF(w_i, d) \) is the number of times word \( w_i \) occurs in document \( d \), \(|E|\) denotes the total number of documents in the training set and the \( DF(w_i) \) is the number of messages containing the word \( w_i \) at least one time.

We consider the text in title field and body field of messages separately but discriminatively. Usually, title is the outline of body contents, so words in title filed are more descriptive and discriminative in contrast to the words in body field. So, words in title field are assigned larger weights to reinforce their stronger impact. For \( TF(w_i, d) \), one time appearance in title field equals to \( t \) times appearances in body field. The cosine method is adopted to compute the similarity between the document vector and the concept vector in the knowledge map, and thus the documents belong to the certain concept with the maximum similarity value.

**Message Type Identification**

Diverse content analysis schemes have been proposed to classify transcripts of online asynchronous discussion (Wever et al., 2005), which reflects a wide variety of approaches and differ in their level of detail and the type of analysis categories used. Particularly, researchers (Gunawardena et al., 1997)(Kanuka et al., 1998)(Lally, 2001) argue that learners' discussion comprises a series of phases in terms of collaborative knowledge building (e.g. information sharing and comparing, concept exploring and discovering, and negotiation of meaning and construction of knowledge). Nevertheless, according to large amount of experimental study, most of learners' discussion transcripts actually fall into the first phase of information sharing and comparing that includes observations, opinions, statements of agreements, examples, clarifications, and identifications of problems. Accordingly, we
classify the messages in the forums into six types, which is useful to recognize the continuation of the topic regarding information sharing and comparison. Herein the messages types include Question, Opinion, Suggestion, Recommendation, Request and Citing. Thus, learners can search for messages with the specified type to cater for their individual demand. For example, if a learner wants to obtain some learning resources on the topic “semantic web”, he can input the word “semantic web” and select recommendation type. Question-type messages are the most common ones, and we further classify it into four categories corresponding to the four levels of knowledge, aiming to provide question-initiating messages accumulated from the former learners to help address later learners’ same or similar problems:

- **Concept-level questions.** The concept-level questions are the simplest ones, which mainly concerns the explanation or definition of a concept as well as some attributes of a concept.
- **Axiom-level questions.** This type of question is relevant to some common sense and the relationship between concepts, and the answer is “yes” or “no” with validating documents. Moreover, in some cases, the questions began with “what”, “who”, “name” and “which” interrogatives are relevant to the axiom. Such questions are the most ambiguous when it comes to determining the question type. This type of questions can be formulated as an embedded expression Relation (concept, concept) after being pre-processed.
- **Rules-level questions.** Rules mainly indicate the logical relationship between axioms, and causality relationship is typical, so this type of questions includes the questions began with “why” interrogative and other semantics similar questions that ask the reason for some axioms.
- **Method-level questions.** This kind of question takes into account finding a solution to a given problem, which is mainly related to the how-term and other semantics similar questions such as “The approach of ...”. Obviously, this type of questions is the most complicated and the solutions to them will probably involve the reasoning according to the other three levels of knowledge.

<table>
<thead>
<tr>
<th>Type of Messages</th>
<th>Pattern &amp; Hint-keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
<td>What is</td>
</tr>
<tr>
<td>Concept-level (Q&lt;sup&gt;C&lt;/sup&gt;)</td>
<td>What’s the definition of &lt;noun_description&gt; ?</td>
</tr>
<tr>
<td></td>
<td>Who</td>
</tr>
<tr>
<td>Axiom-level (Q&lt;sup&gt;A&lt;/sup&gt;)</td>
<td>Is &lt;noun_description&gt; a &lt;noun_description&gt;?</td>
</tr>
<tr>
<td></td>
<td>{Do</td>
</tr>
<tr>
<td></td>
<td>Which &lt;noun_description&gt; is &lt;adj_description&gt;?</td>
</tr>
<tr>
<td></td>
<td>Name &lt;noun_description&gt; &lt;clause_description&gt;?</td>
</tr>
<tr>
<td>Rule-level (A&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>Why {can</td>
</tr>
<tr>
<td></td>
<td>Why &lt;verb_description&gt; &lt;noun_description&gt;?</td>
</tr>
<tr>
<td></td>
<td>What’s reason for &lt;clause_description&gt;?</td>
</tr>
<tr>
<td>Method-level (Q&lt;sup&gt;M&lt;/sup&gt;)</td>
<td>How to &lt;verb_description&gt; &lt;phrase_description&gt;?</td>
</tr>
<tr>
<td>Opinion</td>
<td>I think/believe ....; my idea is ....; I agree/disagree ....</td>
</tr>
<tr>
<td>Suggestion</td>
<td>Let us discuss ....; my suggestion is ....; can we + Verb ....</td>
</tr>
<tr>
<td>Recommendation</td>
<td>For download; share; (strong) recommend; (Memo: Attachment, URL address)</td>
</tr>
<tr>
<td>Request</td>
<td>Who can help ....; where can I get ....; can you recommend ...?; (Memo: the name of book, software, paper, or author, etc.)</td>
</tr>
<tr>
<td>Citing (for reference)</td>
<td>cite from ....; reference to ....; (Memo: URL address)</td>
</tr>
</tbody>
</table>

By analyzing a large amount of messages on the forums, we define the most usual patterns and keywords for each type. Table 1 lists some of the patterns and hint-keywords for each type of messages, and more details can refer to (Zhuge et al., 2002). We currently parse each hint-keywords to identify the various types of messages in Chinese. Herein we treat the title distinctively, and we process it by following a three-step process: Chinese word segmentation, hint-keywords matching and pattern matching. The Chinese word segmentation is based on the bilateral maximum matching algorithm, and the pattern matching is based on shallow syntax parsing (Fung et al., 2007). At the same time, we adopt the HowNet (Dong et al. 2007), a large lexical database of Chinese, to determine the synonyms and hypernyms for keywords matching. If the title is too short or there is not any matching hint-keyword or pattern, we further analyze the content of the initiating-message to identify its type. Regarding the case of multiple matching hint-keywords or patterns for one message, some rules will be executed to select the final type.
For example, the type of question or request will be considered the most probable because of their typical distinctiveness.

**Semantic Association of Messages**

A semantic link network (SLN) is a model to intuitively represent the semantic relationships between document fragments or documents (Zhuge, 2003). In this paper, we use the SLN to organize the messages with semantic associations. The semantic associations represent the binary relationships between messages, which is denoted as \( M_i \rightarrow \alpha \rightarrow M_j \), where \( \alpha \) is a type of semantic relationship while \( M_i \) and \( M_j \) are messages. The relationship can be one of the following types:

- **Sequential**, denoted as \( M_i \rightarrow \text{seq} \rightarrow M_j \), which defines that \( M_i \) is the prerequisite to \( M_j \). A single message may have multiple prerequisite messages, and can also be a prerequisite to multiple messages.
- **Part-of**, denoted as \( M_i \rightarrow \text{par} \rightarrow M_j \), which defines that \( M_j \) is semantically a part of \( M_i \).
- **Similar-to**, denoted as \( M_i \rightarrow \text{sim} \rightarrow M_j \), which defines that \( M_j \) is similar to \( M_i \). The similar-to link is intransitive.
- **Reference**, denoted as \( M_i \rightarrow \text{ref} \rightarrow M_j \), which means that \( M_j \) is related to \( M_i \), e.g., \( M_j \) can be an annotation of \( M_i \).
- **Cause-effect**, denoted as \( M_i \rightarrow \text{ce} \rightarrow M_j \), which means that \( M_i \) is the cause of \( M_j \), and the \( M_j \) is the effect of \( M_i \).
- **Contrast**, denoted as \( M_i \rightarrow \text{con} \rightarrow M_j \), which means that \( M_j \) is in contrast to \( M_i \). Unlike the similar-to link, this contrast link emphasizes the differences between \( M_i \) and \( M_j \) in a context. The contrast link is intransitive.
- **Build-on**, denoted as \( M_i \rightarrow \text{bo} \rightarrow M_j \), which means that \( M_j \) serves as the supplementary or additional content to \( M_i \).
- **Rise-above**, denoted as \( M_i \rightarrow \text{ra} \rightarrow M_j \), which means that \( M_j \) is the summary or conclusion to \( M_i \). Actually, a single message is rise-above on multiple messages.
- **Empty**, denoted by \( M_i \rightarrow \emptyset \rightarrow M_j \), which means that \( M_i \) and \( M_j \) are semantically unrelated.
- **Null**, denoted by \( M_i \rightarrow N \rightarrow M_j \), which says that no semantic relationship between \( M_i \) and \( M_j \) is known for certain.

Semantic links between messages can be manually defined by forum participants, and automatically be discovered and derived under heuristic rules. The reasoning rules can be used for chaining the semantic relationships and obtaining the reasoning result from the chaining. A simple case of the reasoning is that all the semantic relationships have the same type, which is called single-type reasoning. According to the transitive characteristic of the semantic relationships, we have the following reasoning rule: \( M_i \rightarrow \alpha \rightarrow M_j \), \( M_j \rightarrow \alpha \rightarrow M_k \), \( M_n \rightarrow \alpha \rightarrow M_m \Rightarrow M_i \rightarrow \alpha \rightarrow M_m \), where \( \alpha \in \{ \text{seq, par, ce, ref, bo, ra} \} \). The heuristic rules suitable for connecting different types of semantic links are listed in Table 2.

**Table 2. Reasoning rules for message-SLNs**

<table>
<thead>
<tr>
<th>No.</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule1</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{par} \rightarrow M_k \Rightarrow M_i \rightarrow \text{par} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule2</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{ref} \rightarrow M_k \Rightarrow M_i \rightarrow \text{ref} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule3</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{bo} \rightarrow M_k \Rightarrow M_i \rightarrow \text{bo} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule4</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{ce} \rightarrow M_k \Rightarrow M_i \rightarrow \text{ce} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule5</td>
<td>( M_i \rightarrow \text{ce} \rightarrow M_j, M_j \rightarrow \text{par} \rightarrow M_k \Rightarrow M_i \rightarrow \text{ce} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule6</td>
<td>( M_i \rightarrow \text{bo} \rightarrow M_j, M_j \rightarrow \text{par} \rightarrow M_k \Rightarrow M_i \rightarrow \text{bo} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule7</td>
<td>( M_i \rightarrow \text{bo} \rightarrow M_j, M_j \rightarrow \text{ra} \rightarrow M_k \Rightarrow M_i \rightarrow \text{ra} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule8</td>
<td>( M_i \rightarrow \text{seq} \rightarrow M_j, M_j \rightarrow \text{bo} \rightarrow M_k \Rightarrow M_i \rightarrow \text{seq} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule9</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{ra} \rightarrow M_k \Rightarrow M_i \rightarrow \text{ra} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule10</td>
<td>( M_i \rightarrow \text{par} \rightarrow M_j, M_j \rightarrow \text{con} \rightarrow M_k \Rightarrow M_i \rightarrow \text{con} \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule11</td>
<td>( M_i \rightarrow N \rightarrow M_j, M_j \rightarrow \alpha \rightarrow M_k \Rightarrow M_i \rightarrow N \rightarrow M_k )</td>
</tr>
<tr>
<td>Rule12</td>
<td>( M_i \rightarrow \emptyset \rightarrow M_j, M_j \rightarrow \alpha \rightarrow M_k \Rightarrow M_i \rightarrow N \rightarrow M_k )</td>
</tr>
</tbody>
</table>
After recognizing the topics and types of messages, the transcripts are structured based on the heuristic rules by referring to the knowledge map. Figure 2 intuitively shows a part of Message-SLN. As it illustrates, the ellipse represents a topic-SLN, the square represents a threaded message, and the single-line arrow with label represents various relationships between messages within a topic-SLN. In this sample, there are only three kinds of relationship: reference, build-on, similar-to. The double-line arrow with label represents relationships between different topic-SLNs.

**Figure 2. An example of Message-SLN**

**Topic-centric Social Context Measurement by Discovering SIGs**

Social context is the identical or similar social positions and social roles as a whole that influence the individuals of a group. One important aspect of an individual’s social context is the people with whom the person interacts. A given social context is likely to create a feeling of solidarity amongst its members, who are more likely to keep together, trust and help one another (http://en.wikipedia.org/wiki/Social_environment). Therefore, by computing peer similarity based on topic association and browsing activities, this paper aims to discover special interest groups for supporting participants with friendly social context. Special interest group brings together persons focusing on the same interests, and they communicate at large to address concerns and raise awareness about some problems. They are relatively tightly organized, providing opportunities for sharing expertise and first-hand knowledge on the latest development trends while fostering and prompting the progress in the specific field. So, the kernel to discover SIGs within discussion forums is to detect the learners’ relationship as well as their interests according to their delivered messages.

**Combining SNA with Semantic Relations for Constructing SIGs**

Social Network Analysis relies usually on homogenous networks with one type of actors, namely persons, which are also called one-mode-networks. In such networks, centrality, prestige of a person or the overall centralization and density of a network can be computed by well-defined formulas. This kind of network utterly depends on the reply-to-message relation in the forum and neglects the potential semantic relationship between the topics. Yet computer-supported interaction and cooperation typically involves mediating artifacts, such as written messages within discussion forums. These “tangible” products can be seen as another type of entity to be included in the network. Exchange on the level of the original artifacts can also be provided by similarity-based search in a shared collection of objects (Hoppe et al., 2005). If we restrict our analysis, in this sense, to “communication through artifacts” (Dix et al., 2004) this will result in networks only with relations spanning between elements of the two different categories (i.e. person-artifact). This network structure corresponds to a bi-partite graph.
Figure 3 illustrates the framework to detect special interest groups. The normal way to analyze the discussion transcripts corpus is to use SNA to count the reply-to relationship between learners, which results in a one-mode network. By adding the topics to which the messages belong, the one-mode network can be easily transformed into bi-partite network. On the other hand, community can define the knowledge map to express the domain knowledge. By building the semantic mapping from the topics in bi-partite network to concepts in KM, a theme-centered network can be constructed to indicate the persons gathered around one concept. In this way, the theme-centered network denotes the potential interests of the persons, and by adding the reply-to relationship, a special interest group can be formed with respect to each concept in the KM. Persons might be interested in other persons related to an object of their interest, to contact them for discussion and exchange.

After recognizing message topics with the above-mentioned method, the bi-partite network is transformed into the thematic-centered network by mapping the threaded messages to the concepts in the knowledge map. In the thematic-centered network, learners are linked to the themes if their delivered messages belong to the theme. To select the learners with centralized interest in one theme to form a special interest group, formula 2 is used to compute the energy of the $i$th learner for the $k$th theme. Only when the learner’s energy is larger than the threshold, can he or she belong to a special interest group.

Energy:

$$E_{ik} = \frac{w_1 \cdot I_{ik} + w_2 \cdot R_{ik}}{N_i}$$

Where $N_i$ denotes the total number of messages posted by the $i$th person, $I_{ik}$ denotes the number of $k$th-theme-related messages initiated by the $i$th person, and $R_{ik}$ denotes the number of $k$th-theme-related replying messages posted by the $i$th person. $w_1$ and $w_2$ are parameters that adjust the relative impact of messages. The threshold is set as $\sum_{i=1}^{M_k} \frac{E_{ik}}{M_k}$, where $M_k$ denotes the total number of learners linking to the $k$th theme in the thematic-centered network.

After discovering special interest groups within discussion forums, the following step is to compute several criteria for a SIG member, including participation, mutuality and activity.

Participation:

$$P_i = \beta \cdot \sqrt{\frac{I_i}{\sum_{j=1}^{M_k} I_j}} + (1-\beta) \cdot \sqrt{\frac{R_i}{\sum_{j=1}^{M_k} R_j}}, \beta \in [0,1]$$

77
Where \( I_i \) denotes the number of messages initiated by the \( i \)th person, \( R_i \) denotes the number of messages replying to others posted by the \( i \)th person, \( l \) denotes the total number of members in a special interest group.

**Mutuality:**

\[
M_{ij} = \frac{S_{i-all} \times S_{j-all}}{S_{i-j} \times S_{j-i}}
\]  \hspace{1cm} (4)

Where \( S_{i-all}, S_{j-all} \) are the number of messages that \( i \) and \( j \) reply to others’ messages and sent to others by means of private short-message, respectively. \( S_{i-j}, S_{j-i} \) is the number of messages that \( i \) (\( j \)) replies to \( j \) (\( i \)) and sent to \( j \) (\( i \)) by means of private short-message. So in the network for a SIG, the shorter the length between the nodes is, the more close the relationship between the persons is, and vice versa.

**Activity:**

\[
AC_i = \frac{N_i}{\Delta t_p} \times \frac{1}{(t - t_d) + \tau}
\]  \hspace{1cm} (5)

Where \( N_i \) represents the number of messages posted by the \( i \)th person during the period \( \Delta t_p \), \( t \) is the current date and \( t_d \) is the date the \( i \)th person posted the latest message in the forum. \( \tau \) is the adjust parameter to avoid the denominator is zero, and it is initially assigned 1.

Once a person belongs to a special interest group, he or she will be informed of other learning companions to enhance their in-depth communication and learning. Also, the discussion board can proactively push the newly emerged messages on the specific theme or other related information to the member by means of email or other instant messenger. Such pushing service guarantees the members can keep track of the timely progress on the theme even if they do not login to the forum.

**Collaborative Knowledge Building for Evolving KM**

It is well known that the forum provide convenience for learners to discuss on some common interesting topics, and this constitute a cycle of personal and social knowledge-building. During the collaborative learning process, some new concepts will emerge while some concepts will vanish. So, we believe, instead of being constructed once for all, the knowledge map should evolve in accordance with the progress of the domain. That is, the knowledge map is dynamically constructed with the contributions from all the participated learners, which reflects the up-to-date knowledge in the specific domain.

Topic detection and tracking (TDT) research (Allan et al., 2003)(Brants et al., 2003)(Kleinberg, 2002)(Zhang et al., 2002) mainly focuses on detecting and tracking events in streaming news data. TDT systems monitor continuously updated news stories and try to detect the first occurrence of a new story; i.e., an event significantly different from those news events seen before. Based on the approaches, the KM is initially constructed by the experts or advanced learners, and then periodically analyzing the messages to recognize the possible new topics. We combine the messages in a discussion thread into a document, and then compare a new document with the past documents and make a decision regarding the novelty of the message based on the content-based similarity values. If the new document is not similar to any past documents, it will be labeled particularly. When such documents amounts to a threshold, a new concept with additional describing terms that represents documents will be added to the KM whilst marked red to show its novelty. Likewise, if no documents are classified into concepts in KM for a period of time, the concepts in KM will be marked grey to denote its inactivity.

**Supporting Intelligent Facilities**

**Semantic Search with Relational Navigation Support**

In contrast to traditional search engines, the query-keywords input can be denoted with certain semantic information, for example a concrete concept in the knowledge map. With this semantic denotation, the query results can be
profoundly-repacked with the inward semantic schema of the concept, including basic information and semantic relationships with other concepts. With the advantage of semantic inference with logic foundation, more implicit information can be extracted, and the search results can be more comprehensive and rational. Additionally, learners can specify the type of information he or she want to find, such as questions, recommendation, etc, which leads to the focused search and quick location of the useful information. If the learner does not limit the type of messages, the resulting messages are listed in the default type sequence. Regarding each type, we take into account several criteria to rank the messages, including the number of clicking, the number of replying, and the authority of the message authors.

Instead of listing search algorithm (see (Li et al., 2006) for more details), we address the search process more clearly by a search example. When a learner proposes a query “ontology mapping” without type constraint, the system returns all the relevant messages. Firstly, process the query to search for the matching concept C in the KM as the anchor node. Secondly, semantic associated information must be extracted. There are two main semantic-associated concepts B and D for the anchor node via relationship of similarity and subtype. Additionally, by using the reasoning rules (e.g. transitivity) on the semantic relationship (e.g. Subtype, Sequential) between concepts, the prerequisite concept A is also selected. In this way, all the chosen concepts with their belonging messages are displayed to the learners in a well-structured and coherent manner. A schematic navigation view is depicted in Figure 4.

Recommendation

We use the following formula to compute the authority of a learner within a special interest group, and thus each learner can be informed of other learning peers with different authority on a common theme.

\[
A_i = \alpha \frac{\sum_{t=1}^{T} \left( \frac{I_i^c}{\max(I^c)} + \frac{I_i'}{\max(I')} \right)}{2I_i} + (1-\alpha) \frac{I_i^g}{I_i}, \alpha \in [0,1]
\]  

(6)

Where \( I_i^c \) denotes the times of clicking by others for the \( t \)th message initiated by the \( i \)th person, \( I_i' \) denotes the times of replying by others for the \( t \)th message initiated by the \( i \)th person. \( I_i^g \) denotes the number of messages delivered by the \( i \)th person that are set as good messages. \( I_i \) denotes the total number of messages delivered by the \( i \)th person. \( \max(I^c) \) and \( \max(I') \) denotes the maximum clicking-times and replying-times for a message, \( I^g \) denotes the total number of good messages in this SIG.

Additionally, messages are organized and displayed in a multi-mode to cater for the different learning needs of learners:

- **Best-messages view**, which allows the learners to browse the messages that are set as good by the host.
- **Popular-messages view**, which enables the learners to browse the messages that draw a lot of attention from others by ranking the messages according to the number of replying and clicking.
• Messages-type view, which empowers the learners to quickly locate the messages with specific types, such as question, opinion, request, or recommendation.

• Peer-centric message view, which enables the learners to know about all the messages delivered by a peer. This mode is helpful for novice to learn from other peers with high authority on the specific theme.

Experimental Study

In order to evaluate the effectiveness of the approach presented in this paper, we performed an experimental study and compared the analysis feedback provided by assessors of a discussion forum with the analysis results obtained by the proposed approach. We used the W3CHINA discussion board (available at http://bbs.w3china.org/index.asp) as the data source. We randomly selected 763 discussion threads in a view of “Semantic Web and ontology”, with a total of 4512 messages from the source. We asked two postgraduate students whose research direction is semantic web to assess each discussion thread. They manually labeled the messages with the belonging concept in KM and the type of messages. We then compared these manual labeling results with the labels assigned by the proposed automated approach.

![Image: The defined partial Knowledge Map](image)

Figure 5. The defined partial Knowledge Map

Figure 5 shows the partial knowledge map defined by the organizer of the forum, and the messages are automatically classified to be instances of the concepts. We use the following two formulas to compute the micro-average-precision and micro-average-recall for the classification approach, where $a$ denotes the number of auto-labeling belonging messages indeed belonging to the class, $b$ denotes the number of auto-labeling belonging messages not belonging to the class, and $c$ denotes the number of auto-labeling not belonging messages indeed belonging to the class. Table 3 illustrates the evaluation results for the automated classification.

$$\bar{r} = \frac{\sum_{c \in C} a}{\sum_{c \in C} a + \sum_{c \in C} c} \times 100\% , \quad \bar{p} = \frac{\sum_{c \in C} a}{\sum_{c \in C} a + \sum_{c \in C} b} \times 100\%$$

(7)

<table>
<thead>
<tr>
<th>Table 3. Evaluation for the automated classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>User-labeling belonging</td>
</tr>
<tr>
<td>Auto-labeling belonging</td>
</tr>
<tr>
<td>Auto-labeling not belonging</td>
</tr>
<tr>
<td>Micro-average evaluation</td>
</tr>
</tbody>
</table>

On the other hand, we use formula 8 to compute the success-rate of message-type auto-identification approach, where $P_{manu}$ denotes the number of messages annotated by the assessors, and $P_{auto}$ denotes the number of messages annotated automatically. Table 4 shows the messages-type identification comparison for the manual-labeling and auto-labeling, from which we can see that the overall success-rate for identifying message types is relatively
satisfied. But there are 14 messages that cannot be classified into any class. After analyzing such messages, we found that the main reason is that the title and content of the initiating messages don’t contain any typical hint-keywords or match any templates.

Table 4. Messages-type identification comparison

<table>
<thead>
<tr>
<th>Message type</th>
<th>Approach</th>
<th>Manual labeling</th>
<th>Auto labeling</th>
<th>Common</th>
<th>Success-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td></td>
<td>363</td>
<td>354</td>
<td>340</td>
<td>90.1%</td>
</tr>
<tr>
<td>Opinion</td>
<td></td>
<td>185</td>
<td>192</td>
<td>174</td>
<td>84.8%</td>
</tr>
<tr>
<td>Suggestion</td>
<td></td>
<td>30</td>
<td>33</td>
<td>28</td>
<td>90%</td>
</tr>
<tr>
<td>Recommendation</td>
<td></td>
<td>81</td>
<td>84</td>
<td>79</td>
<td>91.8%</td>
</tr>
<tr>
<td>Request</td>
<td></td>
<td>80</td>
<td>89</td>
<td>78</td>
<td>85.7%</td>
</tr>
<tr>
<td>Citing</td>
<td></td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>91.6%</td>
</tr>
<tr>
<td>others</td>
<td></td>
<td>14</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\[
\text{success} – \text{rate} = \frac{P_{\text{manual}}}{} \cap P_{\text{auto}} \times 100\%
\] (8)

We have implemented an integrated toolkit VINCA with comprehensive functionalities to support content analysis with intelligence and visualization features (Li et al., 2007). It can be installed stand-alone or support online downloading of the discussion transcripts from forum (currently support KF (http://kf.cite.hku.hk) and W3China forum) to conduct analysis. Figure 6 shows the interface to set parameter and conduct text analysis, which provides foundation for executing the message theme-classification and type-identification, and accordingly learners’ semantic relationship matrix can be obtained (see figure 7). Afterwards, in order to intuitively illustrate the discovered SIGs, we imported the obtained relationship matrix to the SNA analysis software UCINET to draw the networks for SIGs on different themes.

Figure 8 and figure 10 respectively show the original constructed relationship network on “DL” and “ontology”, in which numerous learners are associated as long as they delivered messages on the themes. Comparatively, figure 9 and figure 11 show the discovered SIGs on “DL” and “ontology” with learners’ energy filtering, respectively. As figure 9 and figure 11 illustrate, the number of learners decrease greatly and the organization structure of the SIGs become more clear-cut and visible. Taking the SIG on “DL” as an example, we can see that the learner 3771 (ID number) plays the most important role in this SIG, who draw a lot of attention from others. Learner1, 4440, 1, 11958, 9173, 22294 play a secondary role, and learner 1 has more close relationship with learner 3771. Especially, there is an isolate learner 16413 who has no any association with others.
Figure 7. The interface to export learners’ relation matrix

Figure 8. Original relationship network on “DL”

Figure 9. Discovered SIG on “DL”
Figure 10. Original relationship network on “ontology”

Figure 11. Discovered SIG on “ontology”

Figure 12. Structured search results augmented with related themes
The prototype of modules has been implemented towards the semantic forum. Figure 12 gives a snapshot of the search module interface within the discussion forum. It illustrates the search results for the query “ontology mapping”. As the figure shows, a learner can select one item from the drop-down list to specify the type of information that he or she intends to search, or search for all types of information by skipping this selection. As a result, the matching messages are classified and displayed in a structured manner. Each message is described with type, title, author, focus (times of clicking and replying), date and replier. To augment the search results, related themes are also recommended at the bottom of the interface supporting learners’ extensive and further navigation. Thus, a learner can easily click the related themes to get more focused information. Figure 13 shows the recommendation interface for users to browse the messages in a multi-mode, such as best messages, popular messages, etc. Furthermore, more messages posted by learning peers are also recommended for extended learning.

![Figure 13. The interface of multi-mode recommendation](image)

**Discussion**

The popularity of online discussion forums enables learners to discuss problems concerning with the conception, development, implementation, etc. in some fields flexibly and freely. But because of the discrete discussion topics, response-freeness and uncertain partners, this kind of loosely organized mode hinders the learners to conduct in-depth communication and cooperation. If the learners could constitute some special interest groups with relatively steady and tight relationship, they will have more motivation and enthusiasm to involve in the deeper communication and interchange of ideas and knowledge.

On the other hand, online discussion forums usually only support simple navigation by listing messages in sequence of date, author or focus. But as the messages increasingly grow over time, learners feel difficult to navigate through the large amount of messages to find the useful information they really need. Currently, the search functions provided on forums are quite weak. Simple keyword-matching-based search leads to the large returning results and some related information missing, so learners still have to spend a lot of time browsing the messages for useful messages and they are easily to get lost during the navigation. Therefore, in order to accommodate the increasingly added messages within the forum, it is indispensable to improve the search precision and coverage so as to alleviate the navigation burden for the learners. Especially, along with semantic-associated information, search results are displayed in a well-structured and coherent manner. In this way, semantic forums offer great convenience and flexibility for learners to support their informal learning and problem-solving.

**Conclusion**

This paper presents a semantic forum for both active knowledge-rich interactions and peer-to-peer collaboration. Its kernel idea is to organize the discussion transcripts in a semantic link network rather than in a discrete and incoherent
structure, and to discover special interest groups with indicators as authority, participation, mutuality, and activity, allowing learners to closely collaborate with peers as they acquire and build the evolved knowledge repository. In addition, the proposed semantic search empowers the learners to conveniently and effectively access the useful information with relational navigation support. The experimental results show that the approach is feasible and efficient, enabling the effective discovering of interest groups and proper demand-driven navigational guidance.

Ongoing work is to enhance the approach performance by considering the context in a thread, to apply the forum in the practical learning settings and improve it according to the feedback from learners.

Acknowledgment

This research work is supported by National Natural Science Foundation of China (NSFC: 60705023). All data source has been kindly provided by the W3CHINA organization in China. We would like to thank Mr. Xu Han for his cooperative work on data analysis.

References


ABSTRACT
We describe our work on seeding and building a developmental learning community for recruiting and retaining women in educational programs and careers related to computing and information science (CIS). Our approach is explicitly participatory: we work with young women to better understand how they view and approach CIS and to build activities that may attract women who would otherwise be disinclined toward computing. In this paper we focus primarily on what we have learned through focus groups with women at different developmental levels with respect to CIS, and the tools we have been building to engage them in community-building activities.

Keywords
Online community, Learning community, Developmental community, Women in information sciences

Introduction

As girls enter their teenage years, they often lose interest in the areas related to computer and information science and engineering. By the time teenage girls are ready for college their interest in computing and information science (CIS) majors falls far below that of their male counterparts (Edwards et al., 1997). One consequence is that the number of women graduating with CIS-related degrees has dropped by almost 25% in the past ten years (Leonard, 2003). This decrease bodes poorly for the number of qualified CIS professionals, but even more so for the diversity within this profession. Many factors influence young women’s lack of interest, including social stereotypes that computer programming is for “nerds,” the competitive culture of university courses that involve programming, and little encouragement for young women to enter into computing as a career (AAUW, 2000; Barker & Garvin-Doxas, 2004). The research reported in this paper studying the role of learning communities as an approach to counteracting this trend.

Researchers and educators are exploring a variety of interventions to counteract this tendency among young women to avoid CIS fields. For example, both men and women are attracted to pair programming techniques involving cooperative problem-solving (Mather, 1997). Universities often sponsor female student groups to provide peer support and mentoring for classmates (Blum, 2001). Another common outreach approach is to provide summer enrichment experiences to middle or high school girls (Gorriz & Medina, 2000; Kyng, 1995; Pedhazur, 1997). We are building on these efforts in the wConnect project, a learning community in which high school girls, undergraduate women, and female faculty and alumnae come together for collaborative activities that help the women to explore and learn about concepts, skills, and practices related to CIS education and careers.

Education scholars over the past two decades have defined learning communities as groups of learners who organize themselves to collaborate on authentic tasks, describing, explaining, listening to, and interpreting each other’s ideas (Lave & Wenger, 1991). Learning communities often structure their learning by scaffolding that is embedded in the activities and the tools of the community (Bruner, 1960). Learners develop by participating in their community’s discourse, where they encounter and contribute to situated negotiation and re-negotiation of meaning.

In the present project, we are studying developmental learning communities. A developmental learning community organizes its activities into developmental phases, and its members take on corresponding roles and responsibilities for developing themselves and other community members (Rosson & Carroll, 2006). The learning is developmental in the sense that members successively traverse phases and roles, and by joining the community a member makes an implicit commitment to support this community objective.

Members of developmental learning communities understand – either implicitly or explicitly – the phases that correspond to development in their activity domain. These phases may involve skill mastery (as in a martial arts community), and transitions may be formally acknowledged as progress is made (e.g. “apprentice”, “practitioner”,...
“master”). Such transitions often occur when a member meets a community standard or practice, for example a skill test, a body of knowledge or experiences evaluated by other members, a prescribed level of insight that is expressed by the member, or an episode that demonstrates progress. In the wConnect community, the levels include female high school students, some of whom are oriented to CIS and some not; female college students, who also vary in their experience and commitment to CIS as a career; and professional women who have made career commitments but vary in their level and breadth of experience.

Along with an understanding of different developmental levels, community members share an understanding of the roles that are appropriate to each level. To some extent this would be true across any multi-level community, for example a general role of “advice giver” that is accepted by those who are more expert in some community-relevant knowledge or skill. In our community, this role emerges naturally and easily within the university population, for example when women who are in their third and fourth year have already “declared” their intended major and area of specialization offer tips to the first and second year students. However, other roles are more specific to the community, for example when an alumna shares opportunities for internships or jobs that would be a good match for the interests of a college woman, or provides custom advice for “presenting” oneself in an interview for her company. The undergraduate women are also able to take on an “CIS instructor” role toward high school girls, a less likely role for someone who is still taking her own coursework in college.

Finally, members of developmental learning communities share a motivation to advance their own and others’ development. Although this motivation is likely to be more pronounced for women who are long term and active members of the community (e.g., through the social cognitive processes of self-perception and self-consistency, see Riding & Raynor, 2001), the simple act of joining the community reinforces an interest and orientation toward personal growth. Social ties among members reinforce this orientation; social ties lead members to care about others in the community, so that they are more willing to expend effort to expand the community and help each other make progress. A developmental community may also provide rewards for its to promote co-members’ learning, such as increased social capital or more explicit forms of recognition.

Our starting vision for wConnect as a developmental learning community is summarized by the diagram in Figure 1. Our expectation is that female undergraduates will form the core of the community: advanced undergraduates will reach down to their less experienced cohorts (including women not yet decided about their field of study); at the same time first and second year students will reach back to their former high school contexts, inviting high school girls to join them in hands-on workshops and online. Finally, as the college students complete their education and transition into careers, they will reach back to the university students, sharing with them an array of experiences and advice relevant to professional work.

Figure 1. The multi-level structure of the wConnect developmental community
Although considerable attention has been paid to the needs of learning communities in general, little if any concern has been directed to the characteristics and requirements of developmental learning communities. We recognize that many learning communities have developmental goals as part of their missions, but in \textit{wConnect} these goals are the focus. We are taking a long term participatory development approach (Carroll et al., 2000), in which we contact and interact with prospective members at different ages with the dual purpose of learning about their current activities and interests and recruiting them as members of \textit{wConnect}. In the balance of the paper we first describe our participatory methods and themes that we have identified, then turn to the online sites or “places” we are building in support of the multi-faceted \textit{wConnect} community.

**Participatory Design of \textit{wConnect}**

The \textit{wConnect} project has adopted a long term participatory design approach for two reasons. First, to fully understand and meet the needs of the women considering and pursuing CIS education or careers, we need their direct and continuous involvement (Carroll et al., 2000). Second, to ensure personal buy-in and sustainability of the resulting community, we need to recruit and activate a broad range of personal engagement and effort (Merkel et al., 2004). We are also influenced by the observations of scholars who have studied the adoption and appropriation of collaborative technology into real world collaborations. These researchers emphasize the importance of building new services on top of \textit{existing practices} rather than inventing novel “software solutions” (Grudin, 1994) and simply expecting that they can be integrated into the workplace.

The selection of design methods has a significant impact on the final shaping of an artifact in use. Recently, there has been an increase in contextual rather than formal methods for the design of information systems. These methods use a bottom-up approach that involves collection and analysis of empirical data directly from the environment. Very often, prototyping is used to verify and visualize user requirements and also to form a basis for discussions by developing a common ground among the designers and users involved in the process (Muller, 2003).

Irestig et al. (2004) compared two contextual methods – user-centered design (UCD) and participatory design (PD) – to determine whether and how their differences might influence the resulting systems. They collected data from design efforts performed in the same organizational environment and having the same goal and compared the resulting prototypes. One finding was that the UCD prototype was relatively normative and idealistic in its concepts, while the PD system was a better fit to the current work environment with all its imperfections and contradictions. For example, the PD design group explicitly deliberated over whether they should design the “best” system or a slightly inferior system with a better chance to be accepted and chose to optimize for acceptability. Moreover, the PD prototype was more scalable than the UCD prototype.

Another consequence of PD processes is that the resulting prototype may be more accessible (and thus modifiable) after the design team has departed. Many researchers have explored this as an approach to more sustainable design interventions (Dittrich, Erikson & Hansson, 2002; Fischer, Ostwald & Seeding, 2002; Merkel, 2004; Rosson et al., 2007). Indeed the general paradigm of \textit{meta-design} is aimed at just this outcome, namely the design of tools that enable domain experts to develop and maintain their own computing applications and data (Fischer & Scharff, 2000).

Beyond these general arguments for participatory design, these processes are critical for understanding and supporting the needs of developmental communities (Rosson & Carroll, 2006). Because the phases associated with development in these communities are at once defining and quite variable, it is important for community members to help in articulating what the phases are and how they should be supported. The roles appropriate to members at different levels may not be known at the start of the community; direct participation and enactment of community behaviors can help these expectations be become recognized and reinforced as the community’s culture develops. Finally, even though joining a developmental community implies motivation toward its goals, members’ personal motivations are rooted in a diverse mix of activity contexts. For example the rewards and costs for community participation will be quite different for college students than professionals, simply because of the huge differences in their socio-organizational contexts. The reward structure for high school students is likely to be quite personal; for college women it may be tied to their education goals; and for professionals it may be a mixture of personal outreach and industry recruiting. A long-term PD process can be very effective in uncovering and responding to stakeholders’ distinct perspectives on motivation and participation (Carroll et al., 1998; Merkel et al., 2004).
Opportunities and Challenges for Community Building

Our primary method for learning about our target population is focus groups. We organize interactive sessions with 2-6 young women at a time, using a semi-structured group interview script to probe experiences and interests with respect to online interaction with friends and acquaintances. In each discussion, we begin by hearing from each participant about her own background with technology, what she uses it for more generally, and how she uses online tools for social interaction. This interaction is used to establish common ground and social rapport with participants. We then shift to more focused discussions of online communication technologies; in this second phase we emphasize different issues for different age groups. For example, when talking to high school girls we want to see if there are online activities that they can imagine as an extension to what they currently do. For the college women, we are more interested in whether and how they might use an online community to address current or future issues relating to their CIS education or careers. For the alumnae, we try to find out whether and how they might contribute to an online community to encourage younger women (e.g., in college) to make the most of their studies and career plans.

To this point we have met with three high school groups, five college, and three alumnae. Each session has lasted from 60-90 minutes, depending on the number of participants and their level of interest. We have taped most of the sessions, but to this point have relied more on the extensive notes that we take while interacting. We turn now to a summary of major themes we have drawn from our discussions thus far.

Table 1. Themes from High School Focus Groups

<table>
<thead>
<tr>
<th>Current online tools</th>
<th>IM, MySpace/Facebook, web browsers, email, wikis or blogs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extra-curricular interests</td>
<td>Volleyball team, TV series, science olympiad, community theatre, marching band and orchestra, theme movie nights</td>
</tr>
<tr>
<td>Personal ideas for data-oriented websites</td>
<td>TV DVD inventory; shopping comparisons; theatre props schedule; band marching scripts; science projects</td>
</tr>
</tbody>
</table>

Table 1 summarizes some high-level results from the high school focus groups. These young women shared many different sorts of computing activities that they already do, but conveyed little interest in creating new tools to support their everyday activities. For them, an online activity should either address a specific need (e.g., collecting information for college), or should just be what they already do (e.g., email, post to blogs). Not surprisingly, the girls reported a broad range of online social interaction in their current lives, including not just IM, email and web browsing, but also more sophisticated activities like a group wiki for staying in touch with peers. They also reported a broad range of interests and extra-curricular activities. At the same time, the girls needed a fair amount of prompting to come up with novel examples of dynamic websites that could be useful in their everyday lives. When shown an example of a dorm room planner, some of the girls were able to propose ideas that had similar functionality for other aspects of their lives, for example managing the inventory of their community theatre company.

In our focus groups with university students, we focused more specifically on the issues for women in CIS, and what might draw them to an online community system in support of these issues. They told us quite clearly they do not want to spend time online worrying about the general challenges of being a woman in a CIS setting. By the time they reach our program, they have heard this “story” too many times to want to spend more effort on it. Rather they are interested in specific tips about getting summer jobs or other employment options. For instance, we were surprised to hear that they are not particularly interested in establishing 1:1 mentoring relationships with older students or professionals, at least partly because mentoring would be perceived an extra personal responsibility that must then be integrated into busy schedules.

In general, all of the students felt that having an online community for themselves and their peers would be useful, and were enthusiastic in general about possible interactions with high school girls and professional women. Some of these students find it difficult to attend face-to-face meetings of clubs created to support undergraduate women in our college, simply because of their busy schedules. Having an online space to meet and chat as an alternative was thus seen as an attractive alternative. At the same time, the women distinguished between online socializing and having online discussions of professional topics like business culture and values, or how to negotiate salary offers or raises. With respect to online socializing, the women are quite satisfied with existing tools (e.g., Facebook, email) but they believe that a more private or protected space would be better for professional topics.
With respect to the design of the wConnect online community, the university participants offered a variety of suggestions about features that could help or hurt participation; these are summarized in Table 2. With respect to positives, there was an emphasis on news, but also on social connections, e.g. access to information about friends, a chance to network with other women. A number of more specific suggestions about system details were also made, for instance that the website should look good, have an easy to use navigation scheme, and so on. Note that although getting “news” seems like a desired feature, there was also concern about not getting too many notifications, especially emails that do not contain any real content, or newsletters that arrive too frequently.

Table 2. Themes from Undergraduate Focus Groups

<table>
<thead>
<tr>
<th>Attractors for online participation</th>
<th>Detractors for online participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Just being part of a major</td>
<td>No regular updates</td>
</tr>
<tr>
<td>Email notifications</td>
<td>Newsletters that are too frequent</td>
</tr>
<tr>
<td>Fresh content (events, job postings)</td>
<td>Too many emails, especially automatic emails that just repeat the same stuff</td>
</tr>
<tr>
<td>Advice on interviews</td>
<td>No responses to my own or others’ postings</td>
</tr>
<tr>
<td>Contact information</td>
<td>Not interactive</td>
</tr>
<tr>
<td>A friend’s profile</td>
<td>Not aesthetically pleasing</td>
</tr>
<tr>
<td>Networking with other women</td>
<td>Difficult to navigate</td>
</tr>
<tr>
<td>Easy navigation; user friendly</td>
<td>Too many ads</td>
</tr>
<tr>
<td>Aesthetic appeal</td>
<td></td>
</tr>
<tr>
<td>Clickable headlines</td>
<td></td>
</tr>
<tr>
<td>Interactive like a Wiki</td>
<td></td>
</tr>
</tbody>
</table>

Our interactions with female alumnae have been grounded on the premise that there are too few women in the CIS workforce and that there might be ways for professionals like themselves to address this imbalance. We also wanted to understand whether and how wConnect membership might impact their current work context, for instance could it create a strengthen a recruiting channel, or add to their status as leaders in their workplace. Like the discussions with undergraduates, we have organized these interactions into brainstorming sessions about what online activities that might make participation in an outreach effort more fun, easier to manage and make them feel more connected and engaged as alumnae. The general reaction was again quite positive, with most participants believing it would be helpful in attracting and retaining women in CIS. However, two participants from a company with a strong image and associated expectations for employees’ behavior told us that they may not be able to get involved without permission, because their management is always concerned about online public activities. A secure and authenticated environment would make participation easier for them.

Table 3. Themes from IST Alumnae

<table>
<thead>
<tr>
<th>Attractors for online participation</th>
<th>Detractors for online participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing awareness about their companies.</td>
<td>A one-on-one relationship may be too involved on either side.</td>
</tr>
<tr>
<td>Students would like to hear about the kinds of projects they do in their company.</td>
<td>Students may think that alumnae are responsible for decision on internships/jobs.</td>
</tr>
<tr>
<td>Students could learn about the different kinds of job they could have.</td>
<td>Leaving behind virtual footprints.</td>
</tr>
<tr>
<td>Alumnae could match information about a project at their work with a course that a student likes.</td>
<td>Feeling some conflict of interest.</td>
</tr>
<tr>
<td>Help students write targeted resumes and prepare interviews for their company.</td>
<td>Professional vs. personal identity in social networking.</td>
</tr>
<tr>
<td>Alumnae could get credit from their company for participation on this online community.</td>
<td>Talking about ongoing projects when there is no match to students’ interest may turn off some students.</td>
</tr>
<tr>
<td>Alumnae find which courses attract students to their company so they can target those courses to do a class project.</td>
<td>Lack of anonymity.</td>
</tr>
<tr>
<td>Small bio and advice from alumnae on their profile.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 summarizes some of the specific attractors and detractors for online participation that came out of our discussions with alumnae and that we plan to carry forward in our design of the online community system. Not surprisingly, many of the attractors reflected some of the work-oriented benefits these professional women might gain from participating in the community, for example visibility of their company with students, or more specific
connections between projects they know about at work and the interests or courses that being pursued by younger community members. Many of the detractors reflect general concerns about balancing their involvement in this “outside” activity with their primary focus of succeeding in their jobs.

Exploring a Landscape of Online Places

A central goal of our interactions with women at varying levels of development has been to understand whether and how an online environment might be helpful to them in meeting their personal, social, education, or work goals. Our intention is to create an online environment that will come to be a “place” in the original sense discussed by Harrison and Dourish (1996), that is a virtual location that women can visit, feel a personal connection to, and an appreciation of the sorts of activities and practices the place supports. To a great extent, we are merely using best practices for online community building to attract young women considering CIS education and careers, but of course as the group begins to cohere and evolve, we expect that the content will become much more specialized to the computing field.

![Figure 2. Five online places for different sorts of community activities within wConnect](image)

Because we are attempting to work with rather different segments of the developmental community, and the needs of each segment are different, we have been developing a number of different online places in parallel – that is, as we have recognized a need for them, as we are beginning to understand the requirements for each. Our progress thus far is summarized in Figure 2. In the upper left is a website used by the project team to plan and to collect and organize resources; in the upper right is a more static website designed to communicate the community’s developmental vision, aimed primarily at alumnae and their companies; in the lower left is a group defined in the Facebook social networking tool, which serves as a lightweight entry into the social context of the community; in the lower right are
the online tools used for high school workshops; and in the center is the “core” community site built as an extension to Facebook. Ultimately we hope that all members will join the core community, at least in the sense of registering for the application and completing a community profile, but in the shorter term, we are using other parts of the infrastructure to reach out to different sub-populations in different ways.

A Place for wConnect Planning

In support of the small group of individuals defining and organizing the community-building activities of wConnect (including 4-5 undergraduate women and 1-2 graduate students), we have constructed an open-ended website using our Bridgetools software infrastructure. Bridgetools is an open source Java-based collaborative tool suite that supports both synchronous and asynchronous interaction addressed to a set of shared information objects (see http://bridgetools.sourceforge.net). This planning site can be visited and edited using a wiki-style editor in a web browser, but the same content can also be accessed and edited in a synchronous and component-oriented fashion by multiple users at the same or different locations. This is an authenticated space; all members are female and part of wConnect. Individuals present themselves online in an informal but personally identifiable way (e.g., marybethR).

Using the tools available as part of the Bridgetools suite, members can design new sorts of web pages, integrate database content if and as needed, create a customized calendar, and so on. As a member develops skills and interests, she may emerge as a “software developer”, a Java programmer who can create totally new sorts of shareable objects. This website is also used to document the plans, activities, and outcomes of the workshops developed and delivered by university students to high school girls. Different content can be created with different levels of access (e.g., something that is relatively unfinished may be viewable or editable only by its creator).

This wConnect place is the most flexible and open-ended, and affords members the most active and varied engagement in CIS concepts and skills. However, it is still a relatively “invisible” place within the community; most members do not even know of its existence. Its users are the 4-5 community members who work with us as researchers while also participating as community members. These young women are critical to our participatory development approach, because they are both co-designers and early adopters of the community-building tools and activities.

A Place for wConnect Learning

A central developmental activity that links university students with high school girls is the hands-on workshops that introduce the younger girls to fundamental concepts underlying online information systems. The workshops are developmental for the college women who plan, deliver and evaluate the sessions; they are developmental for the younger girls who are exposed to novel concepts and skills. We have been working for the past two years to create a protocol for these workshops that addresses a number of challenges, including:

- Attract participation by high school girls who are not currently inclined toward CIS education.
- Engage the young women in brief (e.g., an hour or two) “programming” activities that can be led by wConnect members, that the high school girls will experience as meaningful, and that will convey basic concepts and skills in computing.
- Leverage pre-existing social connections to jump-start feelings of community, reaching out to high schools for which current wConnect members are alumnae, and recruiting members of friendship groups.
- Create an online learning “place” that can be instantiated and used reliably in a variety of high school settings, with whatever equipment and facilities are available.
- Persuade the high school participants to begin a longer-term relationship to wConnect, so that they have community identities and a place to visit and interact after the workshop is over.

These are major challenges, and our place for high school activity will continue to evolve as we experiment with more high school settings and populations (see Rosson et al., 2009 for a more detailed report of these activities). However our preliminary experience has been that high school girls can follow and complete a simple web development project that conveys the basic CIS distinction between “data” and “user interface”. The image in the lower right of Figure 2 illustrates this online place, which integrates a specialized collaborative workspace tool (built
using Bridgetools components) with a standard web browser for viewing and interacting with the resulting application.

During the high school workshops, girls work alone or in pairs on a small web project (for example, a web system that can support the college application process). When they leave, their applications continue to be available to them, either for actual use, or for continued development (in the workspace). At the same time, the participants are invited to join the larger wConnect community. But regardless of whether the girls enlist as active members of wConnect, we create a social link with them, and may contact them at intervals as they move through high school.

A Place to Share the wConnect Vision

A third element in the wConnect landscape is a place that is entirely open, designed for public access. As information-oriented website, this element has less “placeness” than the other components. However it does play a critical role in the overall project scheme. For example, we learned through interactions with alumnae that they would like to communicate the ideas of wConnect to their management and peers, and more specifically have help in obtaining institutional support for involvement. Although many alumnae are willing to participate on a personal basis, they recognize that their involvement will be more valuable if they represent their company. The public website is one vehicle for them to share the community’s goals and activities with others. The website documents key elements of the vision, community members, activities and so on. It also provides links to brochures and posters in PDF form so that community members can conveniently advertise aspects of the project to others they believe may be interested.

Alumnae or other potential sponsors (e.g., some other association aimed at increasing female participation in IT) may be directed to this page by a current community member, or they may simply discover it via a general search, or through a link on another portal. If they wish to know more, there is contact information as well as a brief description of who we are hoping to recruit as community members, and what the benefits are for different classes of stakeholders.

Because this website is our “public image” it is not likely to change as rapidly as other community places, and we are also more concerned with its look and feel. A primary goal is to offer a simple but professional view of the community and its activities with the hope of attracting new members or other outside support.

A Place for wConnect Socializing

As we reported in Table 2, the college women distinguished between discussion of professional topics and everyday social interaction. Most of them assumed that social exchange would occur in the “place” that they already use for such activities, namely the Facebook social networking site. Thus an early “community building” steps was to create a wConnect group in Facebook. To create a sense of privacy, the group is closed - women must submit a request to join. This may happen if they learn about the group from a member, receive an explicit invitation to a community activity, or discover it in other ways while online.

The wConnect group in Facebook is intended to be very lightweight, with little if any community-building activities other than the simple posting and discussions that are built into Facebook. Members can post messages or other information, share and comment on photos, and so on – just as they do in their other social networking places. We view this online place as a low-commitment starting point for new members who may be uncertain of wConnect’s objectives and whether or how the community can be helpful or fun for them, but are willing to identify with the general idea of female solidarity in CIS.

A Place for wConnect Community-Building

Our current focus is on creating a place that will “grow” with the community. We had at one point hypothesized that the Bridgetools website would evolve to serve this purpose, but its unfamiliar tools and user interaction techniques (e.g., relative to Facebook or other popular sites) seem to get in the way. According to our participant-researchers,
these tools feel as if they are for research not community interaction. In contrast, we have had good success with the group concept in Facebook, with about 50 members thus far. As a result we have begun to design a community application that uses the public API from Facebook. It overlaps somewhat with what wConnect members can do in Facebook (e.g., hold online chats, participate in a forum, share photos, etc.), but it is different in that members must first officially join the community. We are hoping to find a balance between the familiar and fun activities young women are used to pursuing in Facebook and related online places, and the private feeling of having a place of their own to discuss more sensitive career and development issues.

Member profiles in the wConnect application are tailored to the individuals’ education or professional interests and are maintained in a community database, not in Facebook. The same is true for discussions or chats hosted by the community. We are currently exploring a range of other more novel activities as well, for example a map through which members can leave notes about places they have lived, worked, or visited, and a network visualization that will let members find their connections to other members based on community-relevant attributes (e.g., career interests, courses taken, tools used, etc.). Member have also proposed to add games that could serve both to introduce members to one another and offer a background of fun or relaxing activities during members’ free time. Several members have stepped up to the role of “Facebook developer” so that they can take a more direct role in supporting new activities like these within the community site. Given our commitment to participatory development, we expect that a wide range of activities will emerge as the community evolves.

Current Status of the wConnect Community

The wConnect project has been underway for almost two years, but we are still exploring techniques for recruiting new members, especially at the high school and alumnae levels. We have used a variety of methods to recruit members from different levels, and have had some success with each, including a) surveys distributed within many undergraduate classes in our college, with an option to indicate interest; b) visits and brief talks to a college organization for female undergraduates; c) email and slides shared in college classes, timed to be just before an online meeting of wConnect; d) focus group invitations emailed to female alumni through contacts maintained by the college development office, timed to occur with recruiting events; e) workshops delivered by college members to students at their former high schools; and workshops delivered as part of a larger university outreach program to high school girls; and personal invitations by members to other women who they know and think might be interested. One of our current projects is to document and organize these different approaches so that they can be disseminated to other organizations interested in building similar communities.

Through these mechanisms we have obtained contact information (name, email address) for over 150 young women who have either explicitly joined wConnect or expressed interest in doing so at some point. Most of our efforts to follow up with interested individuals have been focused on high school and college students, but we are now beginning to reach out to alumnae. For example, we have contact information for about 40 alumnae who have indicated a general interest in our project but who have not yet emerged as active members of the online community. We also have not yet reached out broadly to female graduate students or faculty in our own college, because we expect that these individuals will be much easier to engage and we wish to first establish a core community of active undergraduate members.

Although the recruitment and enrollment of wConnect members is ongoing, we are finally achieving critical mass. Our new Facebook application supports a community of about 50 members. The application hosts profile information, discussions, photos, newsletters, and chats. Although the community extends the Facebook API, the data it maintains and presents is stored independently, so that members need not worry that their wConnect activities might be observed or monitored by Facebook or its many users. As hoped, the online group is diverse with respect to developmental levels, with 12 high school students, 27 undergraduates, 4 graduates, 5 alumnae and 2 female faculty.

Challenges and Future Directions

A significant challenge for wConnect is to strike the right balance of effort and reward for women to join and participate in the community (Grudin, 1994). We know from our interviews that most of the women we seek to
engage are already busy with many other personal, education, or career-oriented activities. We hope that our multiple places approach (i.e., affording multiple options for involvement) will promote broad participation.

Another challenge has been to offer community development tools that are simple and “fun” enough to promote community use, but also powerful enough to engage new CIS knowledge and skills. Our approach to this challenge is twofold: we are extending the Facebook application to attract more activities across developmental levels, including those at the high school age. At the same time, we are investigating ways to go beyond the Facebook platform, because we have found that Facebook is less commonly used by the advanced undergraduates as they prepare for the workforce, or by the alumnae who are now young professional. We are also exploring techniques for supporting the high school hands-on workshops inside Facebook, so that their end-user programming activities can operate on our own community information. We hope to move to a situation where interested high school girls or undergraduates could explore and acquire CIS knowledge and skills independently of the college-led workshops.

This short paper has offered a snapshot of the work we have done thus far in initiating and supporting an online developmental community for women interested in CIS education and careers. However we are far from providing a summative view of whether and how well our community building techniques will succeed. By design wConnect will form and grow according to the interests and energies of its members. This makes its difficult to predict its trajectory. As researchers, our job is to ensure that the activities and tools we provide are engaging, useful and usable, but to let the members discover their own developmental paths.

Acknowledgements

This project is supported by The National Science Foundation, CNS-0634337. Thanks to Gretel Sheasley, Valerie Werth, Anastassia Iojanina, Lauren Hayward, Tim Payone, Nick Oja, Sam Lolla, Tisha Hansford and Danielle Nitopi for their help with different aspects of the project.

References


Emerging Web Technologies in Higher Education: A Case of Incorporating Blogs, Podcasts and Social Bookmarks in a Web Programming Course based on Students’ Learning Styles and Technology Preferences

Nauman Saeed¹, Yun Yang¹ and Suku Sinnappan²

¹Faculty of Information and Communication Technologies, Swinburne University of Technology, Melbourne, Australia // Tel: +61 3 9214 4751 // Fax: +61 3 9819 0823 // nsaeed@swin.edu.au // yyang@swin.edu.au
²Faculty of Higher Education, Swinburne University of Technology, Melbourne, Australia // Tel: +61 3 9215 7912 // ssinnappan@swin.edu.au

ABSTRACT

The adoption level of emerging web technologies is on the rise in academic settings. However, a major obstacle in the practice of web-based instruction is the limited understanding of learners’ characteristics and perceptions about technology use. Thus there is a need to understand the relationship between students’ learning styles and their preferences for instructional strategies, including the use of emerging web technologies. Since learning styles provide information about individual differences in learning preferences they can suggest how instruction can be best designed to support the learning preferences. In this paper, a research framework has been proposed to incorporate emerging web technologies into higher education based on students’ learning styles and technology preferences and a case study has been carried out to validate the proposed framework. An action research methodology has been adopted to carry out the study, which comprises of conducting a survey about students’ learning styles and technology preferences; incorporating a combination of emerging web technologies based on the survey findings; and analyzing key achievements and shortcomings of the study to redefine research objectives. The study provides support for the proposed framework by highlighting the significant relationships among students’ learning styles and technology preferences and their impact on academic performance.

Keywords

Emerging web technologies, Learning styles, Technology preferences, Higher education

Introduction

The growth of web-based applications has made the web an important educational medium (Siau, Nah, & Teng, 2002). Notably, with the creation of emerging web technologies such as blogs, wikis, instant messengers (IM), social bookmarks, podcasts and vodcasts, the web is transforming into a fully interactive space and the control of content has been decentralized to allow every one to collaborate, create, publish, subscribe, and share information (Asmus, Bonner, Esterhay, Lechner, & Rentfrow, 2005). However, a major obstacle for the practice of web-based instruction is the limited understanding of learners’ characteristics (Yang & Tsai, 2008). In addition, the success of educational programs involving web-based technologies depends to a considerable extent on students’ acceptance and use of these technologies (Raaij & Schepers, 2008). Therefore students’ learning styles and technology preferences should be taken into account in order to effectively incorporate emerging web technologies into our courses. Due to the recent developments in e-learning applications the education system is in the midst of a fundamental change. This transformation from traditional classrooms to web-based learning environments has changed the learning styles and interactions between instructors and students (Agres, Edberg, & Igbaria, 1998). Therefore it is important to investigate the learning styles and preferences of the current generation in order to effectively incorporate emerging web technologies into their courses.

Since learning styles provide information about individual differences in learning preferences they can suggest how instruction can be best designed to support learning preferences (Akdemir & Koszalka, 2008). A review of learning theory literature suggests that learning styles and preferences influence the effectiveness with which individuals learn. Therefore, a first hand knowledge of students’ learning styles and preferences can help lecturers choosing the right methods of instruction for the right audience (Smith & Dalton, 2005). In this paper, we present a research framework and hypothesize that students’ learning styles influence their preferences for using technology and that the use of appropriate technology positively influence their academic performance. An action research methodology is adopted by conducting a survey to collect students’ learning styles and their preferences for emerging web technologies; incorporating a combination of web technologies in a web programming course based on survey outcomes; and analyzing the effect of above experiments on students’ academic performance. The study outcomes
provide support to our proposed framework which can serve as a guideline for educators who wish to incorporate emerging web technologies into their courses.

**Literature Review**

**Educational uses of emerging web technologies**

The adoption level of emerging web technologies is on the rise in academic settings (Long, 2006; Pulichino, 2006). These technologies include blogs, wikis, instant messaging (IM), social bookmarks, podcasts, vodcasts, etc. A blog (or weblog) is a type of website, usually maintained by an individual with regular entries of commentary, description of events, or other material such as graphics or video (http://en.wikipedia.org/wiki/Blog). A wiki is a collection of web pages that are linked to each other, and reflect the collaborative works of many authors (Beldarrain, 2006). Instant messaging is a form of real-time communication between two or more people based on typed text via Internet (http://en.wikipedia.org/wiki/Instant_messaging). Social bookmarking is a method for Internet users to store, organize, search, and manage bookmarks of web pages on the Internet, typically in the form of tags (http://en.wikipedia.org/wiki/Social_bookmarks). Podcasting is a method of distributing audio recordings via the Internet, allowing users to subscribe to a feed of new files (Walton, Childs, & Blenkinsopp, 2005) while vodcasting refers to the distribution of video files. Several examples of incorporating these technologies into courses can be found (Augar, Raitman, & Zhou, 2004; Belanger, 2005; Farmer & Bartlett-Bragg, 2005; Lu, Chiou, Day, Ong, & Hsu, 2006).

Emerging Web technologies mentioned above have the potential to create engaging learning environments. In academic settings, students and lecturers alike are achieving many of the benefits of these interactions (Baird & Fisher, 2005). For example, blogs facilitate the publication of knowledge, opportunities for subsequent reflection and analysis, and help learners understand the relational and contextual basis of knowledge, knowledge construction and meaning making (Ferdig & Trammell, 2004). Students enjoy blogging and use it for community building, resource-consolidation, sharing ideas or as a personal journal (Kerawalla, Minocha, Kirkup, & Conole, 2008). Wikis facilitate the creation of shared knowledge, dissemination of information, and group interaction (Augar, et al., 2004). Instant messaging promotes collaborative learning and team work, encourages collaborative learning, and promotes interaction and communication skills (Lu, et al., 2006). Social bookmarks allow quick and easy access to online resources and provide an ‘insiders’ guide to information and references (Asmus, et al., 2005). Podcasts and vodcasts provide an innovative and exciting way for people to improve communication, collaboration and social networking (Ractham & Zhang, 2006). They can also be used for dissemination of knowledge, broadcasting news to staff and students, supplementing class materials, guest lecture presentations and as a marketing tool for attracting prospective students (Harris & Park, 2008). The above mentioned features are key learning elements and make emerging web technologies appropriate for educational use. However, to help students achieve full cognitive development, lecturers need to integrate these technologies with the end-user experience and learning styles (Baird & Fisher, 2005). This paper aims at addressing these issues by exploring relationships between students’ learning styles and their preferences of using emerging web technologies in higher education.

**Learning styles and learning preferences**

The cognitive learning theory emphasizes on the importance of individual differences in learning and considers learning style as a criterion for individual differences (Kia, Aliapour, & Ghaderi, 2009). Learning style is defined as signed indicator that how students perceive, interact and response to learning environments (Keefe, 1979). (Kemp, Morrison, & Ross, 1998) found that learning styles form a student’s unique learning preference and help instructors in the planning of teaching / learning environment. (Sadler-Smith, 1996) states that learning style is a distinctive and habitual manner of acquiring knowledge, skills or attitudes through study or experience while learning preference is favoring of one particular mode of teaching over another. Several classifications of learning styles can be found in the learning theory literature including: Hill’s Cognitive Style Mapping, Kolb’s Learning Styles, Dunn and Dunn Learning Styles, Grasha-Reichman Learning Styles, and Gregorc Learning Styles (Ayersman & Minden, 1995). More recently, the Felder learning model focuses on the aspects of learning styles that are particularly significant in engineering and IT education (Zywno & Waalen, 2002). The index of learning styles (ILS) questionnaire developed by Felder and Soloman measures four learning style dimensions: active–reflective, sensing–intuitive, visual–verbal, and sequential–global (Felder & Soloman, 1993). Felder suggests that students learn in a variety of ways. Some tend
to focus on facts, data and algorithms; others feel more comfortable with theories and mathematical models. Some conceive more from visual information like pictures, diagrams and simulations; others get more from spoken and written information. Some prefer interactive learning; others learn well individually (Felder, 1996).

Understanding the relationship among learning styles and instructional strategies holds great promise for enhancing learner perceptions of their own learning (Claxton & Murrell, 1987). (Akkoyunlu & Soylu, 2008) emphasized on the importance of knowing students’ learning styles to design and manage different web-based environments or other learning materials in various subject areas. Several existing studies showed that matching learning styles with teaching methods is advantageous to academic achievements (Huey Wen Chou & Wang, 1999; Lipsky, 1989; Smith & Dalton, 2005). For example, (H. W. Chou, 2001) found clear differences in the performance and learning preferences of field dependent and field independent students in their study of comparing learning styles with training methods. (Sun, Lin, & Yu, 2008), using Kolb’s learning style inventory, also reported that ‘accommodators’ made the most significant ’achievements’ in the study. (Becker, Kehoe, & Tennent, 2007) reported that students preferred traditional methods of delivery as compared to flexible web-based delivery methods. Similarly, (Butler & Pinto-Zipp, 2006) compared the learning styles with online teaching preferences (using Gregorc Learning Styles Delineator) and reported ‘dual’ learning style (Concrete-Random / Abstract-Sequential) as dominant with strong preferences for asynchronous interactions. However, their study did not report any relationships between learning styles and the technologies used. We would like to explore such relationships in our study using Felder’s dimensions of learning styles. Some previous studies also attempted to highlight the impact of learning styles on academic performance and reported that learners with particular learning styles performed better than others. We would also like to explore these trends in our study.

Research Hypotheses

As mentioned earlier that matching learning styles with instructional methods is advantageous to academic achievements. Our research framework is based on the fact that learners’ individual characteristics influence their preferences for using technology and that the use of appropriate technology positively influences the academic performance. We present our research framework in Figure 1 and hypothesize as follows:

- **H1**: Preference/use of instructional technology is influenced by student’s learning style.
- **H2**: Academic performance is positively influenced by the use of appropriate instructional technology.

![Figure 1. Research framework](image)

We would attempt to validate our first hypothesis by examining the impact of students’ learning styles on their preferences for using various technologies. Based on this, we would incorporate a combination of emerging web technologies in course delivery. Such a technology use would address the needs of a wide spectrum of learning styles and help achieving well-balanced performances across all learner types, thus validating our second hypothesis.

Research Methodology

Experimental design

An action research approach was adopted to carry out the study. Action research is considered as a collaborative activity for finding ways to improve teaching and increase student achievement. It is characterized by spiraling cycles of problem identification, systematic data collection, reflection, analysis, data-driven action taken, and finally, problem redefinition (Kemmis & McTaggart, 1988). We followed this approach and identified the following phases in our study:

- To collect students’ learning styles and technology preferences for emerging web technologies.
• To experiment a combination of emerging web technologies based on students’ learning styles and technology preferences.
• To analyze the impact of above experiments on students’ academic performances.
• To identify key achievements and shortcomings of the study and redefine our research objectives.

Participants

The study was conducted during 2007 with a total of 204 students in a web programming course. The participants were studying toward the completion of Bachelors or Masters of IT degrees.

Measurement

The data for this study were gathered by means of online surveys. Learning styles data was gathered using Felder-Soloman’s learning style inventory (LSI) (Felder & Soloman, 1993). As stated earlier that Felder learning model focuses on the aspects of learning styles that are particularly significant in engineering and IT education (Zywno & Waalen, 2002) and is considered as one of the mostly used models to capture individual differences during the last decade (Dag & Gecer, 2009), therefore deemed suitable for our study. Felder’s model classifies students as: active–reflective; sensing–intuitive; visual–verbal; and sequential–global learners (Felder, 1996). The Felder’s LSI consists of 44 questions each carrying two responses (a or b). It provides the scores (as 11A, 9A, 7A, 5A, 3A, 1A, 1B, 3B, 5B, 7B, 9B, 11B) for each of the four scales. Scores 1-3 on either side of the scales represent well-balanced or ‘mild’ preference, scores 5-7 represent ‘moderate’ and scores 9-11 represent ‘strong’ preference - a total of 12 possible outcomes on each scale.

To collect technologies preference data, the authors designed a questionnaire presented in Appendix A. The participants were asked to rate their technology preferences against various academic activities on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Results of learning style and technology preference surveys

Demographics

With a total number of 204 enrolled students, 119 responded to learning styles survey, including 101 (84.9%) males and 18 (15.1%) females, with a response rate of 58.3%. The majority (85.7%) was aged between 21 and 29. For technology preference survey, we received 105 responses including 90 (85.7%) males and 15 (14.3%) females, with a response rate of 51.4%. Nearly 70% reported their Internet usage as more than 15 hours per week, and 82.2% described ‘Study’ as the major reason for using Internet, which suggests that the majority was well aware of the web usage in academia and was familiar to some extent with the web-based e-learning technologies.

Learning styles frequency distribution

*Table 1* presents the mean and standard deviation values for each of the learning style scales. These figures show that the majority of students exhibited well-balanced learning style except those on the visual-verbal scale where verbal dimension was dominant.

<table>
<thead>
<tr>
<th>LS scales</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active – Reflective</td>
<td>.73</td>
<td>4.288</td>
</tr>
<tr>
<td>Sensing – Intuitive</td>
<td>2.45</td>
<td>4.459</td>
</tr>
<tr>
<td>Visual – Verbal</td>
<td>5.8</td>
<td>4.063</td>
</tr>
<tr>
<td>Sequential – Global</td>
<td>2.16</td>
<td>3.959</td>
</tr>
</tbody>
</table>
Relationships within learning styles

The Pearson correlation coefficient revealed two significant relationships within the four learning style scales. A small positive relationship between active-reflective and visual-verbal scales ($r=.191, N=119; p<0.05$); and another small positive relationship between sensing-intuitive and sequential-global scales ($r=.216, N=119; p<0.05$) (In large samples ($N>100$), very small correlations may be statistically significant (Pallant, 2005)). These results suggest that reflective learners were correlated with verbal learners while intuitive with global learners. The strength of above relationships was measured as follows: $r = .10$ to $.29 \rightarrow$ small; $r = .30$ to $.49 \rightarrow$ moderate; $r = .50$ to $1.0 \rightarrow$ strong (Pallant, 2005).

Relationships within technology preferences

In the technology preference survey, participants were asked to rate their technology preferences against various academic activities. Besides collecting the preferences for emerging web technologies we also asked our participants to rate their preferences for some conventional technologies like email and Blackboard as these were used as an integral part of course delivery in our University. Pearson correlation coefficient revealed several significant relationships as shown in Table 2: three strong (figures in bold), three moderate (figures in italic) and two small relationships, all significant at $p<.05$. All relationships appeared positive, i.e., the higher preference for one technology lead to the higher preference for the other. These results suggest that the learning preferences of various technologies were closely related, i.e., our students preferred to try various technologies in their study routines instead of relying on one particular tool. These results were also encouraging for us to experiment a variety of technologies in our study. However, it was interesting to see that none of the emerging web technologies (blog, wiki, IM, podcast, vodcast) correlated with the conventional technologies (email and Blackboard). It should be noted that “Social bookmarks” is not included in Table 2 as it does not correspond to any academic activity mentioned in the technology preference questionnaire (Appendix A).

<table>
<thead>
<tr>
<th></th>
<th>Blackboard</th>
<th>Blog</th>
<th>Email</th>
<th>IM</th>
<th>Podcast</th>
<th>Vodcast</th>
<th>Wiki</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackboard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blog</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Email</td>
<td></td>
<td>.38</td>
<td></td>
<td></td>
<td>.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td></td>
<td></td>
<td>.38</td>
<td></td>
<td>.56</td>
<td>.39</td>
<td>.38</td>
</tr>
<tr>
<td>Podcast</td>
<td></td>
<td>.29</td>
<td></td>
<td></td>
<td>.56</td>
<td>.57</td>
<td>.29</td>
</tr>
<tr>
<td>Vodcast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiki</td>
<td></td>
<td>.38</td>
<td></td>
<td></td>
<td>.57</td>
<td></td>
<td>.29</td>
</tr>
</tbody>
</table>

Relationships between learning styles and technology preferences

To see the impact of students’ learning styles on their preferences of using emerging web technologies, we combined the data from learning styles questionnaire and technology preference questionnaire. The corresponding results suggest that active-reflective scale did not yield any significant relationship when correlated with various technologies. However, sensing-intuitive scale was negatively correlated with the preference of using email ($r=-.12; N=89; p<0.05$) while positively correlated with that of blog ($r=.12; N=89; p<0.05$). This implies that sensors preferred to use email while intuitive learners preferred to use blog. The visual-verbal and sequential-global scales found to be negatively correlated with the preference of using vodcast ($r=-.16; N=89; p<0.05$) and podcast ($r=-.14; N=89; p<0.05$) respectively. These results suggest that visuals preferred vodcast while sequentials preferred podcast. We will discuss more on these results later in the Discussion section.

Matching learning styles with technology preferences

Based on the above statistical analyses, we incorporated a combination of web technologies in the web programming course based on the following rationale: As mentioned in the previous section that podcast, vodcast, email, and blog were the preferred technologies for sequential, visual, sensing and intuitive learners respectively. Therefore, to facilitate these four learner types, we decided to incorporate a course blog, audio (podcast) / video (vodcast) recordings of lectures and email communication in the above course. To match the appropriate technologies for the
remaining four learner types we made some assumptions based on the statistical results presented in the previous sections.

Table 3. Technology preferences for various learner types

<table>
<thead>
<tr>
<th>Learning style</th>
<th>Technology preference</th>
<th>Learning style</th>
<th>Technology Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Social Bookmarks</td>
<td>Reflective</td>
<td>Podcast</td>
</tr>
<tr>
<td>Sensing</td>
<td>Email</td>
<td>Intuitive</td>
<td>Blog</td>
</tr>
<tr>
<td>Visual</td>
<td>Vodcast</td>
<td>Verbal</td>
<td>Podcast</td>
</tr>
<tr>
<td>Sequential</td>
<td>Podcast</td>
<td>Global</td>
<td>Blog</td>
</tr>
</tbody>
</table>

The first assumption was based on the fact that like intuitive learners, global learners would also prefer to use blogs because these two learner types were significantly correlated. Second, we assumed that audio recordings (podcast) of lectures would suit verbal learners since they are characterized as those who learn best by using words, written or oral (Felder, 1996). Following this argument we assumed that like verbals, reflective learners would also prefer podcast because these two learner types were significantly correlated. Finally, to facilitate active learners, who usually prefer new and innovative ways of learning, we assumed that social bookmarks would be a good match for them as social bookmarking is relatively a new trend in academia. Table 3 presents a summary of matching various technologies with each of the learner types.

Experiments with incorporated technologies

Based on the criteria discussed above, we incorporated a course blog, a social bookmarks page and audio / video (podcast / vodcast) recordings of lectures in the web programming course. In order to avoid the hassle of setting up extra blogging and bookmarking software on the University servers, we relied on free hosting services. For the course blog we chose Blogger from Google (http://www.blogger.com), as Blogger is considered fast, reliable and well-known blogging service (Bryant, 2006). To set up the course bookmarks page, we chose Delicious (http://del.icio.us), a well-known social bookmarking service. For the recording and delivery of lecture podcasts / vodcasts, we used the local recording services available at our University. The experiments were carried out for the whole semester from February to May 2007.

Figure 2. Students’ response to weekly ‘minute paper’ and peer mentoring on the course blog

Course blog

The course blog was set up on Blogger with the following aims: (1) to get quick feedback on teaching and learning issues from students; and (2) to enhance student collaboration. To achieve the first objective, we used a weekly “minute paper” on the course blog. Minute papers are considered a way to promote meta-cognition thinking amongst students and to provide academics with un-graded and immediate feedback in order to assess how well and how
much students have learned (Murphy & Wolff, 2005). A minute paper is comprised of two simple questions: ‘what is (are) the most significant thing(s) you have learned in this lecture’, and ‘what question(s) remain unclear in your mind?’ (Figure 2). Students were encouraged to submit their course related problems and queries using weekly minute papers. To achieve the second objective, we encouraged students to post solutions to the queries raised by fellow students. This was intended to help students publish their thoughts, become critical thinkers, help others understand course contents, enhance online communication and collaboration, and create a virtual community, thus improving understanding level of the whole course.

Course bookmarks page

The course bookmarks page was set up on Delicious with the following aims: (1) to build a repository of online (web programming) resources with easy access to every class members; and (2) to engage students in the knowledge building and sharing process.

![Figure 3. Students’ contribution of web programming resources on the course bookmarks page](image)

To achieve the first objective, lecturers posted their online resources (bookmarks) on a weekly basis, covering main topics of the week. For easy access to relevant course topics, the bookmarks were also bundled (grouped) into individual text book chapters. To achieve the second objective, students were strongly encouraged to maintain their own Delicious accounts and post their favorite bookmarks on weekly course topics (Figure 3). Students’ accounts

![Figure 4. Audio/video (podcast/vodcast) recordings of lectures](image)
were then added to the network list, which allowed lecturers to monitor students’ participation. It also allowed students to see what fellow students were contributing.

**Lecture podcasts**

As mentioned earlier, we used local recording services at our University to record and publish the audio and video recordings of all lectures. Students could access the recordings in three different formats: live streaming, manual download, and automatic subscription (*Figure 4*). They were also able to access those recordings through *Blackboard* ([http://www.blackboard.com](http://www.blackboard.com)), which is used as an integral course management system in our University. All lectures were recorded and published with the aim of complementing teaching and learning through podcasting.

**Analyses on academic performance**

At the end of the semester after all assessments were completed for the *web programming* course, we conducted some further analyses in order to see the impact of gender, age and learning styles on students’ academic performance. The Chi square analysis of academic performance and gender variables revealed no significant differences in the overall performance of males and females. This means that we found no evidence in our study that males performed better than females or vice versa. Similarly, no differences were found in the overall performance of young and mature-age students in the class.

The Chi square analysis of four learning style scales and high achievers (who scored 85% or more) across various assessment components (programming assignments, multiple choice questionnaire, and final examination) also revealed no significant differences. For example, the proportion of active learners who scored 85% or more in the overall assessment was not significantly different to that of reflective learners. Similar results were obtained for the low or average performers in the class. These results suggest that our teaching approach, including the use of emerging web technologies, was not biased toward any particular learning style(s). In other words, it accommodated all learner types to achieve well-balanced academic performance. This is consistent with the findings of (Akkoyunlu & Soylu, 2008) and (Sun, et al., 2008) but in contrast with several previous studies such as: (Chamillard & Karolick, 1999), (Thomas, Ratcliffe, Woodbury, & Jarman, 2002) and (Allert, 2004), which were highly biased toward particular learning styles.

**Discussion**

The learning style survey helped in analyzing the dominant learning style, frequency distribution, and correlations within various learning styles. Although the verbal style emerged as the most frequent single learning style, the majority appeared to be ‘well-balanced’, which is consistent with the findings of (Alfonseca, Carro, Martin, Ortigosa, & Paredes, 2006) and (Zywno & Waalen, 2002). (Butler & Pinto-Zipp, 2006) also reported “verbal” as the most frequent single learning style and “dual” (combination of two) learning style as the dominant, in their study of comparing learning styles with the preference of online teaching methods. Our findings also support Butler’s idea that today’s learners are more flexible in stretching their learning styles to accommodate a variety of teaching methods. The correlations within four learning style scales revealed significant relationships between *reflective* and *verbal* as well as *intuitive* and *global* learners, which are consistent with the findings of Alfonseca et al. This is a significant outcome of our study and has implications for forming appropriate groupings in programming assignments or projects.

The relationships within various technology preferences revealed that our students preferred to use both synchronous and asynchronous communication tools in their academic communication. This is in contrast with Butler’s findings which reported the preference of asynchronous tools only (Butler & Pinto-Zipp, 2006). These results also suggest that today’s students are ready to experience new technologies in their study routines and are willing to collaborate using multiple communication channels.
The correlations between learning styles and technology preferences also revealed some interesting outcomes. Sensing learners preferred email more than intuitive learners. Felder defines sensors as those who feel more comfortable with details and happen to be more careful (Felder & Solomon, 1993). Our finding also highlights this characteristic of sensors as they preferred to rely on a traditional communication tool like email as compared to IM, blog, or wiki. Intuitive learners, on the other hand, are defined as those who prefer discovering possibilities and relationships and are always ready to try out new things. This characteristic of ‘intuitors’ is highlighted by their preference for blogs over other traditional technologies like email or Blackboard. The relationship between visual learners and vodcast was a natural phenomenon as ‘visuals’ learn best from pictures, diagrams, flow charts, and demonstrations (Felder & Solomon, 1993). These results also highlight the acquaintance of intuitive and visual learners with the new technology. Similar trends were reported by (Zywno & Waalen, 2002) where active, intuitive and visual learners were the most frequent users of hypermedia-based instructions. Finally, sequential learners tend to gain understanding in linear steps and follow logical stepwise paths, hence podcast was a good choice for them to run sequence of lectures at their own pace over and over again to get a better understanding of the course contents. Above results support our hypothesis H1, which states that instructional technology is influenced by learners’ learning style. The above outcomes can serve as a guideline for lecturers for choosing the right technology for the right audience in their courses.

Another significant finding of our study is the achievement of well-balanced academic performances across all learner types. No significant differences were found on the proportion of high, average or low achievers across various learner types, age-groups, or genders. This is consistent with the findings of (Akkoyunlu & Soylu, 2008) where no significant differences were found between students’ achievement level according to their learning styles. (Sun, et al., 2008) also reported that the web-based virtual learning environment was suitable for various learner types as no significant differences were found in their grade achievements. These results fully support our hypothesis H2, which states that academic performance is positively influenced by the use of appropriate instructional technology. However, this is in contrast with some previous studies, where academic performance was heavily biased toward particular learning styles. For example: (Chamillard & Karolick, 1999) reported that reflective and verbal learners performed better than others in their study. Similar findings were reported by (Thomas, et al., 2002) and (Allert, 2004). At this stage, we can not rule out other contributing factors behind these well-balanced performances such as students’ background of web programming, motivation for the course, assessments, sample size, etc. Also, involving students from a web programming course could have introduced a bias toward a certain type of learning style as most of the students involved would have some kind of scientific background.

Conclusion and future work

The research framework presented in this paper is supported by the data collected through user surveys, experiments performed with incorporated technologies and student performances in a web programming course. We have been successful in highlighting the technology preference of various learner types as well as incorporating a novel combination of emerging web technologies in the course delivery. Our findings suggest that today’s learners are flexible in stretching their learning styles and are able to accommodate varying instructional strategies including the use of emerging web technologies. They further suggest that learning styles of today’s learners are flexible enough to experience varying technologies and their technology preferences are not limited to a particular tool. Our experiments with incorporated technologies and achievement of well-balanced academic performances across all learner types motivate us to extend the study to other courses combined with varied technologies. The inclusion of students from non-scientific backgrounds in our future studies would further help understand the relationships among various learner types and their technology preferences.

References


Appendix A

Learning styles questionnaire (Index of Learning Styles)*

The complete questionnaire is available at:  &lt;http://www.engr.ncsu.edu/learningstyles/ilsweb.html&gt;.


Technology preference questionnaire

SECTION I: Demographics

1. What is your age _______ (in years)?
2. What is your gender?
   a) male
   b) female
3. How many hours do you use the Internet per week?
   a) 5-10 hours
   b) 10-15
   c) 15-20
   d) more than 20 hours
4. For what purpose do you use Internet mostly?
   a) study
   b) work
   c) personal
   d) entertainment
   e) others

SECTION II: Technology preferences

Please read through the following statements and decide how much you prefer to use each of these technologies to perform the following academic activities. Using the scale provided write the number that best indicates how you feel on the space provided before each statement (please enter only digits from 1 to 5).

Least preferred 1 2 3 4 5 Most preferred

1. I would prefer to revise lecture online by:
   a) _____ listening to podcast (audio) of the lecture
   b) _____ watching vodcast (video) of the lecture
   c) _____ discussing it on the course blog
   d) _____ discussing it on the course wiki
   e) _____ discussing it with my peers on IM
   f) _____ discussing it with my peers through email
   g) _____ discussing it with my peers through Blackboard

2. I would prefer to submit online a group project to lecturer:
   a) _____ as a podcast presentation
   b) _____ as a vodcast presentation
   c) _____ through course blog
   d) _____ through course wiki
   e) _____ through IM
   f) _____ through email
   g) _____ through Blackboard

3. I would prefer to have online class discussion with lecturer through:
   a) _____ podcast
   b) _____ vodcast
   c) _____ course blog
   d) _____ course wiki
   e) _____ IM
   f) _____ email
   g) _____ Blackboard

4. I would prefer to have online group discussion through:
   a) _____ podcast
   b) _____ vodcast
   c) _____ course blog

5. I would prefer to have online study discussion with a friend through:
   a) _____ podcast
   b) _____ vodcast
   c) _____ course blog
   d) _____ course wiki
   e) _____ IM
   f) _____ email
   g) _____ Blackboard

6. I would prefer my lecturer to conduct virtual office hours through:
   a) _____ podcast
   b) _____ vodcast
   c) _____ course blog
   d) _____ course wiki
   e) _____ IM
   f) _____ email
   g) _____ Blackboard

7. I would prefer to receive assignments online from lecturer through:
   a) _____ podcast
   b) _____ vodcast
   c) _____ course blog
   d) _____ course wiki
   e) _____ IM
   f) _____ email
   g) _____ Blackboard
Adaptive Role Playing Games: An Immersive Approach for Problem Based Learning

Pilar Sancho, Pablo Moreno-Ger, Rubén Fuentes-Fernández and Baltasar Fernández-Manjón

Department of Software Engineering and Artificial Intelligence, Universidad Complutense de Madrid, Spain // pilar@sip.ucm.es // pablom@fdi.ucm.es // ruben@fdi.ucm.es // balta@fdi.ucm.es

ABSTRACT

In this paper we present a general framework, called NUCLEO, for the application of socio-constructive educational approaches in higher education. The underlying pedagogical approach relies on an adaptation model in order to improve group dynamics, as this has been identified as one of the key features in the success of collaborative learning scenarios. Students’ learning strategies are analyzed considering a simplified version of Vermunt’s model for learning styles. The resulting profiles provide the basis to group students in teams where each student is assigned a role according to his/her learning strategies. The result is the formation of complementary and semi-autonomous learning teams that collaborate to achieve solutions to the problems provided by the instructor. The framework is instantiated through an online multiplayer role-playing game environment, which sets a stage for the underlying collaborative problem-based learning approach. The framework has been benchmarked in different programming courses at the Complutense University of Madrid during 2007-08 and 2008-09 academic years. This work also presents the results of these experiences.

Keywords

Educational games, Learning Objects model, Learning Objects meta-data, Learning Management System interoperability

Introduction

The Net Generation (Tapscott, 1998) has already arrived at university and college. They differ completely from the people in charge of their education (their teachers and parents) in the role that ubiquitous technologies have played in their everyday lives. Today, students have grown up using devices like computers, mobile phones, and video consoles for almost every activity, from studying and work to entertainment and communication. This has probably altered the way they perceive and interact with the environment, both physically and socially (Prensky, 2001b). To meet the different cognitive requirements of the new generation, the educational community is considering new ways of learning. In particular, there is a wide interest in trying to engage students with the appealing features of videogames and Internet tools (Prensky, 2001a).

Nevertheless, most of these approaches have been developed without taking into consideration the current infrastructure of Learning Management Systems (LMS). LMS are probably the most extended tools for managing the whole educational process in higher educational institutions, from teaching and learning to administration. We think this omission is a mistake: even though learning through games has very positive educational values, it obviously cannot cover the whole range of educational needs. Game-related initiatives are often disconnected from key aspects of an integrated educational process such as storing and managing learning content, performing specific learning management tasks or maintaining historical student records.

NUCLEO is an e-learning framework that comprehends a pedagogical strategy and the technological infrastructure to support it. The learning strategy is deeply grounded in the socio-constructive pedagogical stream (Vygotsky, 1978; Palincsar, 1998): it combines Problem Based Learning (PBL) (Savery & Duffy, 1996) and Computer Supported Collaborative Learning (CSCL) (Koschmann, 1994) in a framework that uses a multiplayer role-playing videogame as the delivery format. In NUCLEO, active collaborative learning takes place in the scenario of a virtual world with game-like mechanics. NUCLEO takes the learners, who are represented by avatars, into a fictional scenario where they have to solve a number of missions. To succeed, they need to collaborate with other students within a team. The missions proposed to the teams respect the style and context of the videogame metaphor and are rendered in the virtual scenario in which the activity takes place. The specific fictional scenario and underlying narrative can be changed to fit different subjects and contexts. In addition, the framework implements a plug-in to work along with a LMS, enhancing both tools with complementary functionalities.
One of the key aspects in collaborative learning is the implicit assumption that the participants learn from each other inside teams. Therefore, the way in which students are grouped has a strong impact on the results of the learning process. A positive learning experience might turn into a negative one depending on the group composition. The coordination for group work is another key aspect for the success of a collaborative learning experience. An effective way to coordinate group work is to assign concrete responsibilities to individuals using functional roles (Strijbos, 2004). NUCLEO addresses these aspects by means of an adaptation model that relies on Vermunt’s conception and classification of learning styles (Vermunt, 1992). The “Inventory of Learning Styles” (ILS) proposed by Vermunt is used to try to identify those students who need more intensive guidance through the learning process, and those who are more inclined to drive their own learning experience. By grouping heterogeneous students according to their learning style, we presume that the most autonomous students will provide leadership and guidance to the group. At the same time, the effectiveness of the collaboration process within a team will improve by teaming up students with complementary learning strategies and assigning them concrete interdependent responsibilities linked to a role. Teams are also a way to enrich social interaction among students.

NUCLEO is a complex framework that relies on several hypotheses that need to be experimentally proved. In particular:

- Vermunt’s ILS is an effective underlying model for forming teams.
- Using functional roles helps to improve group work coordination.
- The 3D immersive scenario is positive for motivation.
- The role game dynamic is positive for the motivation and induces students to adopt a more active role.
- The competitive atmosphere created by using social recognition strategies is positive for motivation.
- The framework is effective in terms of knowledge acquisition.
- The framework is effective in terms of helping students to acquire soft skills and team work abilities.

In order to prove these hypotheses in a cost-effective way, NUCLEO is being developed according to a fast prototyping development plan. So far, two different prototypes have been developed, Mundo NUCLEO and Mare Monstrum, each of them conceived to prove different aspects of the underlying assumptions. These two prototypes have already been benchmarked in several cases of study. NUCLEO is a long term research project that, so far, has been running for the last three years. Some results concerning these cases of study have already been published. In particular, (Sancho, Fuentes, Gómez-Martín, & Fernández-Manjón, 2009) discusses the results of the first three cases of study, performed with the Mundo NUCLEO prototype. The main goal of this work was to see if the role game dynamic was effective in terms of increasing motivation; and, also, if students were able to acquire the required technical knowledge included in the course curriculum. In (Sancho-Thomas, Fuentes-Fernández, & Fernández-Manjón, 2009), the results obtained in one case of study with the Mare Monstrum prototype are presented. This case study was conceived to prove if the 3D immersive scenario was positive for motivation and if the framework was effective in terms of helping students to acquire soft skills and team work abilities.

This paper is focused on the new results obtained concerning the effectiveness of using Vermunt’s ILS as the underlying model for the adaptation process. In particular, we discuss its effectiveness in terms of team formation and its suitability for role distribution as a way to improve group work coordination. The results have been gathered through the cases of study performed so far with the two prototypes (i.e., Mundo NUCLEO and Mare Monstrum) when applied to university programming courses.

The rest of the paper further discusses the elements outlined in this introduction with the following structure. First, we describe the pedagogical foundations underlying NUCLEO. Then, we present the overall structure of the framework and describe two NUCLEO-specific instances, with different narratives and game metaphors, already applied in the prototypes of the cases of study. After that, we describe the cases of study undertaken so far, presenting and discussing the results obtained. Finally, we summarize our conclusions and outline future lines of work.

Related Work

The NUCLEO framework combines several existing learning approaches. NUCLEO considers dPBL (distributed Problem Based Learning), CSCL (Computer Supported Collaborative Learning), student centered models using adaptation according to learning styles, and learning in virtual worlds or MUVE (Multi-User Virtual Environments).
These approaches offer a wide range of alternatives for team formation and organization. Our choices have been determined by our specific goals and contexts. There are different works belonging to these three approaches that share some features with the NUCLEO framework. This section presents some of the most similar to NUCLEO in terms of the adaptation model.

Team formation and organization have drawn important attention for both educative and work environments. Literature describes different alternatives according to the goals intended for the team activity (Morgeson, Reider, & Campion, 2005). Since one of the NUCLEO objectives is the acquisition of teamwork skills by students, it considers heterogeneous teams of students. This choice is in the line of some authors (Oakley, Felder, Brent, & Elhajj, 2004) that consider this heterogeneity can enrich the discussion in the group with different perspectives and knowledge, offering students the opportunity to deal with different personalities, and gathering ineffective students with stronger learners, who can provide inspiration and guidance. Nevertheless, this kind of team organization also presents some risks in an educational context (McCracken & Waters, 1999). First, there are personalities that are inherently difficult to work with them. Composing groups with such people can lead to form ineffective team works. Nevertheless, it can be argued that real teams may include that kind of personalities and, of course, no students should be put aside because of their personalities. Second, programming problems with a true need of teamwork require a long time, since they are closer to real projects in the range of months than to toy practices of a few weeks. However, programming courses commonly pose several problems in the course with reduced development times. A heterogeneous team of non-related students can find it difficult to set up the required bonds and routines in such short periods. The problem in both risky situations is that failing to achieve a proper workflow can affect the motivation of the entire team, and prevent the students from acquiring the intended skills. A different interesting approach for CSCL can be found in (Bravo, Redondo, Verdejo & Ortega, 2008).

The other relevant feature of NUCLEO discussed in this paper is the organization of work through functional roles. Industrial software development commonly uses this organization to set up team organization (Brooks, 1975). Thus, NUCLEO considers adopting it as a way to train students in real team work practices, but also to develop suitable soft skills. However, roles are only one of the ways of making students interdependent and therefore setting up their social bonds (Johnson, Johnson, & Smith, 1991). Besides, imposing teams an internal organization that they do not agree can prove to be a futile effort (Oakley et al., 2004). Students forced to work following an imposed organization of work, can externally accept it but internally work in a self-organized way. This misleads the learning goals of the intended approach with roles. NUCLEO tries to avoid this situation through two mechanisms: an imposed composition for the teams and the assignment of different roles; making individuals accountable for the work related with their role.

Finally, there are also implementation aspects of the adaptation model in MUVE and LMS. The integration between them has also been explored through initiatives such as SLOODLE (Kemp & Lingstone, 2006; Gonzalez & Blanco, 2008). However, NUCLEO’s distinguishing feature is that the integration is driven by the underlying pedagogical model described in the following section.

**Pedagogical Model**

The NUCLEO framework proposes an educational model deeply rooted in socio-constructive pedagogical theories (Palincsar, 1998). As previously mentioned, it combines several existing approaches to socio-constructive education, such as CSCL, PBL, student-centered models using adaptation according to learning styles, and MUVE. In fact, NUCLEO’s underlying learning strategy is actually an enhanced PBL approach that makes use of the following strategies in order to reach the pursued educational objectives:

- A role game and a 3D immersive scenario in which students are represented by avatars. This strategy is conceived to address the following educational objectives:
  - To promote active roles between the students, who are incited to abandon their usual passive attitudes (Corti, 2006).
  - To create a propitious atmosphere for the emergence of social and affective bonds among players, leading to the formation of communities of practice (Baron, 1999).
  - To promote competition by social recognition through a game narrative in a 3D scenario. This “good competition” between students can provide an additional push to the enhancement of motivation (Johnson & Johnson, 1994).
An adaptation model based on a simplified model of Vermunt’s ILS (Vermunt, 1992) (see section Adaptation to Learning Styles) with a twofold objective:

- Forming effective semi-autonomous teams in which students are grouped considering their compatibility in terms of their learning strategies using Vermunt’s ILS (Vermunt, 1992).
- Improving group coordination by distributing the work according to functional roles that are also assigned to the students by means of their classification in the ILS. A learning strategy structured by collaboration scripts in order to guarantee students’ participation in educationally meaningful activities.

The rest of this section further describes these pedagogical foundations of our framework.

**Collaboration and Problem-Based Learning**

According to Barrows and Tamblyn (1980), Problem-Based Learning (PBL) can be explained as “the learning that results from the process of working towards the understanding or resolution of a problem.” In most cases, PBL is performed in small groups, fostering discussion and collaborative discovery, as the groups need to work together towards the solution for a specific problem or set of problems. A group has a tutorial leader or facilitator who shares information and guides the group through the learning process. In sum, PBL is a process of building new problem solving skills on prior knowledge by using critical thinking approaches and reflection (Maudsley, 1999). This self-directed and collective approach constitutes a very different way to teach compared to traditional lecture-based approaches.

Educational literature has shown the benefits of using PBL and other approaches that promote active collaborative learning to improve students’ thinking skills (Kreijns, Kirschner, & Jochems, 2003). It has been demonstrated that it leads to deeper levels of learning, critical thinking, shared understanding, and long-term retention of the learning material. Furthermore, collaborative learning also provides opportunities for developing social and communication skills, acquiring positive attitudes towards co-members and learning material, and building social relationships and group cohesion (Johnson & Johnson, 1994).

However, the collaborative nature of PBL is frequently an issue. It is often difficult to arrange the schedules of all the members of a team to participate in work sessions. The possibility of meeting remotely can allow more effective PBL initiatives. New technologies have made possible to delocalize this approach, the so called dPBL (distributed Problem Based Learning) and sometimes CSCL techniques are used as virtual implementations of this pedagogical approach (Resta & Laferrière, 2007).

Unfortunately, it is not easy to implement a PBL approach effectively when we factor out the direct personal contact. This is due to different reasons, and one of the most important ones is that PBL relies very heavily on group dynamics for its success. If group cooperation and cohesiveness are key factors in PBL, we must take into account that the lack of face-to-face interaction affects them negatively. Simply providing students with some remote communication tools does not guarantee the emergence of the social interactions that lead to effective collaboration (Lurey & Raisinghani, 2001).

As it has already been stated, NUCLEO applies two different strategies to address these problems and enhance group dynamics:

- An adaptation model aimed at forming heterogeneous, effective and semi-autonomous teams by means of an adaptation of the classification model of Vermunt’s learning styles.
- A role game and an immersive virtual world in order to promote the affective links that may lead to the formation of communities of practice among players.

**Adaptation to Learning Styles**

User adaptation in e-learning can be characterized as the ability of a system to personalize the learning experience to different individual conditions over time. In general, the adaptation process includes three stages (Brusilovsky & Maybury, 2002): gathering information about the user; processing this information to initialize and update a user...
model; and using that model to provide the adaptive behavior. One of the students’ features that can be considered in these models is their learning styles.

Learning styles have generated a lot of debate over the past few years mainly because, in spite of long empirical efforts to pin them down, their identification remains elusive. Nevertheless, some approaches that describe learning styles as flexible strategies to tackle learning have got positive evaluations in relevant independent studies (Coffield, Moseley, Hall, & Ecclestone, 2004). In particular, Vermunt’s model (Vermunt, 1992) was specially conceived for university students. It is really more a classification of students according to the strategies they usually employ to approach learning than a categorization of learning styles, as they are commonly understood. This view fits with our idea of reaching auto-regulated teams, as it can provide criteria for the organization of teams.

Vermunt classifies students into four types depending on the attitudes they adopt in five different areas of learning by means of its ILS. These four learning styles are: meaning-directed (MD), application-directed (AD), reproduction-directed (RD), and undirected (U). This approach helps to distinguish the students who need more intensive guidance through the learning process from those who are more capable of guiding their own learning experience. Those students who are able to self-regulate their learning processes usually present MD and AD patterns, and they would benefit from a more open teaching strategy. Students who would need stronger teacher control and guidance commonly correspond to the RD and U patterns (see Table 1).

Our team formation strategy is based on grouping three or four students, among which there is at least one MD or one AD student. In every team, students are assigned different roles, embedded in the game metaphor, according to their ILS profiles (see Table 1). The Captain of the crew (assigned to MD or AD profiles) is in charge of project planning and progress monitoring. The Knowledge Integrator –KI- (assigned to RD profiles) is in charge of controlling and supervising that all team members acquire the required knowledge. The member Responsible for Communication –RC- (assigned to U profiles) is in charge of managing communication between team and tutor. All of them have concrete tasks, interdependent responsibilities and specific tools assigned.

<table>
<thead>
<tr>
<th>Role</th>
<th>CAPTAIN</th>
<th>KNOWLEDGE INTEGRATOR</th>
<th>COMUNICATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning style</td>
<td>Meaning-directed (MD)</td>
<td>Application-directed (AD)</td>
<td>Reproduction-directed (RD)</td>
</tr>
<tr>
<td>Cognitive processing</td>
<td>Look for relationships between key concepts/theories; build an overview</td>
<td>Relate topics to everyday experience: look for concrete examples and uses</td>
<td>Select main points to retain</td>
</tr>
<tr>
<td>Learning orientation</td>
<td>Self-improvement and enrichment</td>
<td>Vocational or “real world” outcomes</td>
<td>Prove competency by getting good marks</td>
</tr>
<tr>
<td>Mental model of learning</td>
<td>Dialogue with experts stimulates thinking and engagement with subject through exchange of views</td>
<td>Learn in order to use knowledge</td>
<td>Look for structure in teaching and texts to help take in knowledge and pass examinations. Do not value critical processing or peer discussion.</td>
</tr>
<tr>
<td>Regulation of learning</td>
<td>Self-guided by interest and their own questions; diagnose and correct poor understanding</td>
<td>Think of problems and examples to test understanding, especially of abstract concepts</td>
<td>Use objectives to check understanding; self-test; rehearse</td>
</tr>
</tbody>
</table>
Motivation, Narrative, and Role-Playing Games

As stated in the introduction, authors such as Prensky (2001b) or Tapscott (1998) consider that the Net Generation belongs to a new digital media culture that behaves very differently than their teachers and parents do in profound and fundamental ways. For instance, these students are used to multimedia content connected with hyperlinks, and find the traditional teacher-centered courses quite unattractive, preferring a collaborative learning with peers. Although there have been important advances to make teaching strategies more interactive and appealing to them (Prensky, 2001b), this evolution does not seem to completely fulfill the needs of knowledge disclosure for these students.

Considering these observations, the potential use of videogames as a tool for education has drawn significant attention, especially for PBL approaches. Games present several relevant features that could benefit the students’ learning experience (Gee, 2003). They offer immersive and realistic scenarios with engaging narratives that challenge the user to solve problems embedded in the game. At the same time, they are designed to continually offer problems that are just beyond the limit of expertise of the player. They also deliver the required information on demand and in context, when the player can make a meaningful acquisition of it. Finally, games are often the seed of communities of practice where players spontaneously collaborate to build knowledge to solve the game quizzes. Moreover, in the case of multiplayer games, this collaboration is closer as it happens in the same place and time (Lave & Wenger, 1991).

Despite these positive features, videogames are still a resource of difficult use for education. First, their successful use is not just a matter of changing the presentation style of the traditional contents (Laurel, 2001). This destroys the fun and immersion of the game, and therefore its usefulness as educational resource. It is important to perform a reworking of the teaching material to get its proper integration in the dynamics of the game. Second, videogames are not a universal solution. They need adaptation for a wide range of students’ learning styles and subjects. Third, playing does not produce “per se” a valuable learning experience. This depends on the effort invested in the development of the underlying pedagogical approach. Besides, the results of these environments must be compared against traditional ones to consider if they really bring some kind of improvement in a given educational context.

Considering these potential drawbacks, our proposal adopts a different perspective from the one most frequently adopted in videogames. Instead of trying to disguise the educational content inside a game, we have turned the whole learning scenario into a game, based on the idea that gaming is essentially about solving problems that are immersed in the game narrative.

Figure 1. Two different perspectives of the NUCLEO framework architecture.
Description of the framework: NUCLEO reference architecture

As we have previously mentioned, NUCLEO is a general framework for collaborative PBL that includes a reference architecture to instantiate different applications following its underlying pedagogical principles.

The NUCLEO reference architecture has been designed in a modular way in order to:
- Apply a set of strategies to reach the educational objectives described in the “Pedagogical Model” previously described.
- Be used as a plug-in application over a LMS. This allows the reuse of managing services and tools at the same time that deals with data in a centralized way, which simplifies its integration in the educational infrastructure.

Figure 1 depicts the main components of the NUCLEO reference architecture. The LMS acts as a Persistency Layer and, at the same time, offers an alternative way of accessing the contents without high technological requirements. The Application Layer is in charge of two main duties: to handle the adaptation process (by means of the Adaptation Module) and to coordinate the activity of the students based on the collaboration scripts (by means of the Learning Strategies Manager). The following subsections describe in deeper detail the main components outlined here.

Learning Strategies Manager

Research studies show that when learners are left to their own devices in collaborative learning contexts, they rarely engage in educationally relevant activities (Dillengbourg, 2002). A common practice in CSCL environments is to use collaboration scripts that aim at structuring collaborative learning processes in order to trigger group interactions that rarely take place in free collaboration.

Nowadays, in most CSCL systems, the logic underneath scripts is embedded within the system code. However, there are alternatives to model CSCL scripts in computer interpretable ways that separate the logic from its rendering. In particular, a number of pedagogical modeling languages have been developed for this purpose, such as IMS Learning Design (IMS Global Consortium, 2005), which is considered the de facto standard for pedagogical modeling.

Within our system architecture, the Learning Strategies Manager module is in charge of the creation, storage, and management of the personalized learning strategies and all their components. Currently, these components include phases, roles, activities, environments, tools, and learning objects.

Adaptation Module

This module is in charge of providing adaptive behavior. The adaptation logic of a personalized learning system can be defined in the following terms (Karagiannidis & Sampson, 2004): the constituents (what is being adapted?), the determinants (what is the adaptation based on?) and the rules (the logics that define which constituents are affected by the determinants and how).

The constituents are aspects of the learning experience that are subject to adaptation. In NUCLEO there are two constituents: team formation and individual role assignment. In an indirect way, the learning experience is also personalized as it depends on the individual role. The determinants are the aspects of the learning experience that drive the adaptation. This information is stored in the student model (SM). In NUCLEO the SM is initialized with the results obtained by the students in the Vermunt’s ILS questionnaire. More specifically:
- Team formation. The goal of the first constituent (i.e., teams) is to try to form effective and semi-autonomous teams. The system groups heterogeneous students according to their learning style searching to: a) foster collaboration, as the students interact with peers with compatible learning strategies, and b) to increase group work effectiveness. As previously described this team formation is based on the students answer to the Vermunt’s ILS questionnaire.
- Role assignment. Group performance strongly depends on the handling of increased coordination. As described in the pedagogical model, we use functional roles to afford the work organization and communication between members. Roles appear to be most relevant when a group pursues a shared goal requiring a certain level of task division, coordination and integration of individual activities (Strijbos, 2004).
Figure 2 shows the adaptation cycle for the NUCLEO system. In the first step of the process, the SM is initialized by means of the result obtained in the ILS questionnaire. Then, the groups are formed and the roles are assigned. This automatically leads to a personalization of the learning strategy, as roles imply specific duties and use certain tools of the system in order to fulfill them. The learning cycle is adjusted to the resolution of a mission. After every mission, the SM is updated collecting information from three different sources: the mark obtained by the group in the mission, the individual mark obtained by the student in the peer-to-peer evaluation (every student evaluates his/her teammates) and the frequency of use of the specific role tools. According to this information, students’ roles and teams can be reassigned, for instance, in order to reduce intra-team conflicts or to give students the possibility of experiencing new social contexts.

**Figure 2**. Adaptation cycle for the NUCLEO system

**NUCLEO Graphical User Interface**

NUCLEO uses as Graphical User Interface (GUI) a MUVE that stages the game scenario in a fantastic immersive world in which students are represented by avatars. Digital games are an ideal channel to promote relevant educational aspects such as problem-solving skills or analysis and reflection. Among the educational benefits that videogames can bring to instructional methods, we are very interested in the social bonds formed among players, especially in multiplayer games. These bonds may foster the creation of communities of practice (Lave & Wenger, 1991). The virtual world serves as a setting for an engaging narrative that adds meaning and contextualization to the different missions.
In the current stage of development of the NUCLEO project, two different instances, that use different ambiances and metaphors, have been developed: Mundo NUCLEO and Mare Monstrum. The first instance sets the pedagogical background described above in a fantastic futurist scenario and the second in a medieval one, but both share the same narrative essence. The students belong to a civilization menaced by a terrible enemy and they have to be trained in the weapons of knowledge to confront it. The scenario and the game reproduce a school of warriors in which aspirants have to compete solving different missions that simulate a real attack from the enemy. With this purpose they are integrated in small combat units (usually of 3 or 4 members). Only those who pass the training will graduate as Paladins.

Persistence Layer: LMS

As has already been mentioned, the use of digital games and MUVE as educational tools can provide a number of advantages over traditional education. However, there are some drawbacks that have not been totally addressed yet. For instance, digital games and MUVE are rarely fully integrated in the educational infrastructure and usually behave as “black boxes”. It is therefore impossible to get any instructionally relevant information about the course of the game-based learning experience from them, such as students’ performance for assessments or keep a persistent student historical record.

These issues could be addressed by taking advantage of the already deployed e-learning infrastructure. Many educational organizations are using modern Virtual Learning Environments (also called Learning Management Systems or LMS) not only for distance learning, but also as a complement for traditional lectures within an educational trend usually known as blended learning or b-learning. Those LMS (e.g., Moodle, Blackboard-WebCT or Sakai) (Wainwright, Osterman, Finnerman, & Hill, 2007) are not only content repositories, but rich web-based systems that provide instructors with tools to track and evaluate the performance of the students, keep a record of each student or to promote communication and collaboration between students. Thus, synergies between educational gaming and e-learning could bring together the benefits of both worlds, an approach that the NUCLEO framework actively explores. In fact, NUCLEO is currently integrated with Moodle LMS (see Error! Reference source not found.) that provides the persistence layer and some functionality (e.g. forums and shared repositories). Teachers and students can access both systems.

Application of the framework: cases of study

Up to now, four cases of study have been conducted in two different courses during two academic years, three of them performed in 2007-08 and another one in 2008-09, with the two instances of the NUCLEO framework (i.e., Mundo NUCLEO and Mare Monstrum). The courses chosen for the experiments shared common characteristics such as a decrease in students’ motivation in recent years, resulting in high drop-out rates, low involvement in classes, and low grades.

Our first concern, therefore, was to determine what the effects of the NUCLEO framework were on students’ motivation. Given the difficulties in assessing student attitudes, we adopted the drop-out rate as an objective measure. This drop-out rate is defined as the number of students who do not attend the final exam, as compared to the number of students enrolled in the course.

Even when motivation is a key factor of a successful learning experience, it is not enough to improve the student acquisition of knowledge and skills. In fact, one of the identified common pitfalls in the use of PBL, CSCL, and games in learning is that students focus on the joy of the experience and fail to get the related knowledge (Royle, 2008). In order to determine the contribution of NUCLEO in knowledge acquisition, we considered the results in the final individual exam, which was mandatory for all the students. An improvement in exam grades may imply that more students had been able to achieve the learning goals. The main conclusions drawn from the experiments concerning motivation and knowledge acquisition can be found in (Sancho et al., 2009).

It was also important for our objectives to determine how the students perceived the key aspects of our learning framework (i.e. acquisition of soft-skills (Sancho-Thomas et al., 2009b) and efficiency of the adaptation model). For this paper, we focused on the performance of the adaptation model in order to constitute the teams and the usefulness of functional roles to coordinate the group work. In order to include some abstract notions such as leadership or
commitment in our performance measures (which may be cloaked by a low performance in a specific exam), we also included a peer evaluation mechanism in which the students in each group rated each other’s performance.

The presentation of these results is organized in three subsections. The first one makes a brief introduction of the settings. In the second one, the statistics obtained as result of three cases of study are presented. The third section briefly discusses the global results of the experiments.

**Description of the cases of study**

In this paper we consider the results obtained in three of the four cases of study performed so far, because they all refer to similar programming courses in the Spanish university context. The remaining one was analyzed in a previous study (Sancho et al., 2009) and was applied to a completely different setting (vocational training).

At the University Complutense of Madrid (UCM), in Spain, several schools teach programming courses. The Electrical Engineering School offers the “Programming Fundamentals” (PF) elective course. PF belongs to a five-year degree and it is studied as part of the second cycle (the fourth or fifth year of the degree). It focuses on some programming basics, such as algorithms, program design and coding. The Computer Science School at the UCM also teaches “Laboratory of Programming II” (LP2) in the second year of the three-year Computer Science Technical Engineering degree. This course focuses on object-oriented programming and data structures.

During the 2005-07 period (i.e., over two academic years), the teachers of these courses followed a traditional teacher-centered approach that included lectures in classroom, practical sessions in the laboratories and compulsory final exams. To calculate the students’ final grades, the teachers took into account the marks obtained in the practical sessions (20% of the final mark in LP2 and 40% in PF), and the final exam (the rest of the mark).

During the 2007-08 academic year, both courses implemented the NUCLEO approach (using the Mundo NUCLEO instance) as an optional choice. Sixty PF students and 175 LP2 students participated in this experiment. Given the differences in the scheduling of the courses, the experiment was held in the PF course during the entire semester, while LP2 followed a traditional approach in the first semester and the experiment was performed in the second one. In both courses, students were divided into Experimental (EG) and Control (CG) groups. EG followed the NUCLEO approach, while CG attended traditional lectures. Lecturers organized the NUCLEO students in teams according to their learning profiles, as previously explained. All the teams had a leader who was a self-directed, autonomous student. In the traditional approach, PF students did the class work individually, and LP2 students organized themselves into groups from one to three members.

NUCLEO team formation relies primarily on the assumption that there is a student with good learning practices who plays the role of the captain of the crew. Therefore, in order to evaluate the effectiveness of the ILS as the underlying model behind team formation, we measure whether Vermunt's ILS really identifies strong students accurately by verifying if MD and AD students get better marks than the other profiles. To verify the accuracy of Vermunt's model for assigning the functional roles, we measure students’ satisfaction with their peers, in terms of how their teammates perceive the fulfillment of their responsibilities.

During the 2008-09 academic year, a new experiment has been performed in PF, with some differences from the previous ones:
- All 54 students in the course followed the NUCLEO approach. This time, participation was compulsory.
- We used the Mare Monstrum instance of the framework, instead of Mundo NUCLEO.
- In addition to the information gathered in previous experiments, we also distributed a satisfaction survey aimed at collecting the subjective impression of students about how successful team formation and role assignment were.

**Summary of the statistical results obtained in the three cases of study**

Table 2 contains a summary of the most important results about the adaptation model for each of the courses studied. Students were categorized and assigned to a different role depending on the results they obtained from the initial
learning style questionnaire: MD and AD students were captains, RD students were knowledge integrators and U students were communicators. Every member had to perform different duties inside the team depending on the role he/she was assigned to. At the end of each mission, students were asked to evaluate their teammates, taking into account several issues concerning the fulfillment of their assigned duties. Table 2 reflects the grades obtained in the final exam by the NUCLEO participants according to their role and the average punctuation obtained by each role according to the perception of their teammates.

Table 2. Main experimental results from the 2007-08 and 2008-09 courses for grades obtained in the exam and for peer evaluation results by different profiles

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Students in CG</td>
<td>MUNDO NUCLEO</td>
<td>MUNDO NUCLEO</td>
<td>MARE MONSTRUM</td>
</tr>
<tr>
<td>Students in EG</td>
<td>38</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>Average grade of MD and AD students</td>
<td>6.66</td>
<td>5.87</td>
<td>6.92</td>
</tr>
<tr>
<td>Average grade of RD students</td>
<td>5.36</td>
<td>4.9</td>
<td>5.1</td>
</tr>
<tr>
<td>Average grade of U students</td>
<td>5.53</td>
<td>5.7</td>
<td>5.4</td>
</tr>
<tr>
<td>Average peer-evaluation grade of MD/AD</td>
<td>9.375</td>
<td>9.5</td>
<td>9.65</td>
</tr>
<tr>
<td>Average peer-evaluation grade of RD</td>
<td>7.954</td>
<td>8.2</td>
<td>8.01</td>
</tr>
<tr>
<td>Average peer-evaluation grade of U</td>
<td>8.24</td>
<td>7.68</td>
<td>8.32</td>
</tr>
</tbody>
</table>

Table 3. Students’ Questionnaire and responses

<table>
<thead>
<tr>
<th>Question</th>
<th>% in strong disagreement or disagreement</th>
<th>% in strong agreement or agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I think the use of roles facilitates handling coordination and group work.</td>
<td>39.53%</td>
<td>60.47%</td>
</tr>
<tr>
<td>2. In my opinion, the team assignment process is suitable for the course and works well.</td>
<td>32.56%</td>
<td>67.44%</td>
</tr>
<tr>
<td>3. I think that the virtual tool is useful to coordinate the activities and enhances collaboration.</td>
<td>4.65%</td>
<td>95.35%</td>
</tr>
</tbody>
</table>

At the end of the course, the students completed a survey composed of nine questions, three of which were aimed at figuring out how the adaptation model was perceived by students. These questions and the results obtained are reflected in Figure 3 and Table 3. The answers ranked from 1 (strongly disagree) to 4 (strongly agree).

Discussion of the results

The results in Table 2 show that, for the three cases of study performed during the 2007-09 period with the two instances of the NUCLEO framework, students with MD and AD profiles obtain significantly better grades than the other two profiles (RD and U). This suits our initial hypothesis that MD and AD students have more effective learning strategies than the other two profiles. Nevertheless, no significant difference among the other two profiles can be observed.

In terms of the peer-to-peer evaluation results, MD and AD profiles (who performed the role of captain) also received significantly better marks from their peers than the RD or U profiles. In addition to this, it can be observed for all the cases of study that RD profiles (who performed the role of knowledge integrator) received the poorest evaluation from their teammates. According to these data, two possible conclusions of the accuracy of the model can be extracted: the role of captain is suitable for MD students, while even though the other two obtain fairly good
marks, RD and U students do not perform their assigned duties so well. Knowledge integrator (RD students), in particular, is the less suitable profile for its assigned duties.

For the two cases of study performed during the 2007-08 academic course, participation was voluntary, while in 2008-09 all the students followed the NUCLEO approach. Even though we thought that the volunteer enrollment could have skewed the results of the experiment (with the most enthusiastic students participating in NUCLEO), no significant differences can be observed between both sets of results.

During the 2008-09 case study, we have added an additional source of information to evaluate the system. A satisfaction questionnaire was filled at the end of the course by all the students, asking them to contrast the approach with their previous experience with traditional approaches. Here, it must be noted that PF is only available for students in the fourth and fifth courses. There were three questions related to how the students have perceived the efficiency of the adaptation model.

An overwhelming majority (95.35%) of the students thought that the Mare Monstrum virtual learning environment was satisfactory, with half of them considering it “very” satisfactory. This shows a very high degree of satisfaction with the 3D game environment and its general dynamics.

The percentage is lowered to 67.44% when evaluating the results of the team formation process, with fewer than 10% of students considering it as “very” satisfactory and being the only aspect of NUCLEO that some of them considered “very” unsatisfactory. Two specific reasons were among the most frequently cited for student dissatisfaction. Firstly, although most teams established correct social interactions, some of them seemed to include personalities with a high degree of incompatibility. This incompatibility was related to conflictive mates. These circumstances can probably be avoided by extending the adaptation model to include personality traits focused on social bonds. This kind of study has already been considered in research for team formation (Morgeson et al., 2005).

In any case, NUCLEO has to incorporate mechanisms that articulate conflict resolution to tackle these situations when they emerge in teams. Secondly, a considerable number of students felt uncomfortable working with unknown
mates. They adduced that the team formation process was unsatisfactory because they would have preferred to work with friends. However, given the objectives of the NUCLEO framework (one of which is to help to improve team work abilities and social skills), being able to work with different kinds of people is a key point.

When analyzing role assignment as a way to improve group work coordination, the percentage of satisfied students is 60.47%. In this case, most of the discontented students thought that the only important role within the organization was the captain and that the other two were secondary with marginal tasks. Specifically, students who held the role integrator role were the most discontent with their attributions, as they found them boring and repetitive.

Conclusions and future work

In this paper, we have described the NUCLEO framework to develop adaptive PBL approaches. The framework encompasses a general conceptual model and a reference architecture. Its learning is staged in a 3D immersive world following the mechanics of classical multi-player role based games. This framework has been used to build two different applications (Mundo NUCLEO and Mare Monstrum) that have been benchmarked in the Spanish higher educational context during 2007-08 and 2008-09 academic years.

Even though the results obtained have been globally satisfactory in terms of rising student motivation and final grades (Sancho et al., 2009), we must point out that these conclusions have to be supported with data collected over next few years to get a wider temporal perspective. Only in this way we will be able to know if the approach itself increases the students’ motivation and learning results, or if it is just a side-effect of enthusiasm for the novelty.

Even though the framework is massively perceived by the students to be useful for learning soft and group work skills, and to develop technical knowledge, our results show that some aspects of the system have to be improved. First, the adaptation model to form the teams needs to be enhanced. Vermunt’s model for learning styles has to be complemented with other sorts of information, as personality traits. We are also studying other possibilities for the conceptual model behind the team formation like Kirton’s Adaptation-Innovation theory (Kirton, 2006). In the second place, while the role of captain seems to be satisfactory according to all the measurements of success used (e.g., perception, assigned tasks or linked learning profile), the other two roles have to be reconsidered. The role of knowledge integrator seems to be especially inadequate for RD types.

Acknowledgements

The Spanish Committee of Science and Technology (projects Flexo-TSI-020301-2008-19, TIN2007-68125-C02-01 and TIN2008-06464-C03-01) has partially supported this work, as well as the Complutense University of Madrid (research groups 921340 and 921354), and the EU Alfa project CID (II-0511-A).

References


Collaborative Paper-based Annotation of Lecture Slides

Jürgen Steimle¹, Oliver Brdiczka² and Max Mühlhäuser¹

¹Department of Computer Science, Darmstadt University of Technology, Hochschulstr. 10, 64289 Darmstadt, Germany // steimle@tk.informatik.tu-darmstadt.de // max@tk.informatik.tu-darmstadt.de
²Palo Alto Research Center (PARC), Palo Alto, CA 94304, USA // oliver@parc.com

ABSTRACT

In a study of notetaking in university courses, we found that the large majority of students prefer paper to computer-based media like Tablet PCs for taking notes and making annotations. Based on this finding, we developed CoScribe, a concept and system which supports students in making collaborative handwritten annotations on printed lecture slides. It includes mechanisms for the paper-based sharing and semantic tagging of annotations and slides. Moreover, we present a novel visualization of shared handwritten annotations providing an integrated access to annotations from all learners. The results of a user study indicate that CoScribe efficiently supports student annotation and can be easily integrated into current annotation practice.

Keywords

Annotation, Digital paper, Digital pen, Anoto, Lecture, Computer-supported Collaborative Learning (CSCL)

Introduction

University lectures are often considered as suboptimal learning settings, as the communication is centered around the instructor and learners easily become rather passive listeners. Collaborative notetaking and annotation processes are important elements to overcome these problems and foster active learning (Prince, 2004): Taking personal notes stimulates students to actively follow the lecture, to consciously select important content, and to rephrase it in their own words. Moreover, notes have a reminding effect during review after class (Kiewra, 1989). The sharing of notes with other learners may support students to complete their notes, to critically examine their own understanding, and to co-construct a shared understanding with other learners.

Many electronic systems for student notetaking and sharing have been proposed, e.g. (Davis et al., 1999, Kam et al., 2005, Anderson et al., 2006, Berque, 2006). However, as we found out in a field study, most students do not use electronic tools for notetaking and annotation but prefer traditional pen and paper. Paper can be intuitively annotated, manipulated with both hands, is highly mobile and cheap. Moreover, paper documents can be laid out in physical space for comparing and organizing pages (Sellen and Harper, 2003). Empirical work further indicates that working with paper does not only imply a lower cognitive load than working with traditional computers, but it also has lower cognitive load than working with a digital stylus on a Tablet PC or a graphical tablet (Oviatt et al., 2006). Therefore, combining the advantages of paper with the benefits of digital documents (like electronic search and collaboration) presents novel opportunities for improving classroom annotation, review, and collaboration.

In this paper, we present CoScribe, a concept and system which enables students to collaboratively annotate lecture slides. During class, students use a digital Anoto (http://www.anoto.com) pen to make handwritten annotations on printed slides and empty pages as with a traditional pen. The interaction is therefore intuitive and lightweight. When reviewing the contents after class, a software viewer provides for access to both one’s own annotations and to shared ones of other learners.

This article has three main contributions: 1) Based on two quantitative studies, we present new insights on how students use printed and digital media for taking notes and making annotations in lectures. 2) We discuss the design of paper-based user interfaces for the annotation of course handouts and present two mechanisms for the semantic tagging, indexing, and structuring of printed content. The paper-based user interface is complemented by a novel visualization of shared handwritten annotations for the efficient overview and access to annotations of other users. 3) The design has been validated in a user study, which indicates that CoScribe efficiently supports student annotation and can be easily integrated into current annotation practice.
The remainder of this article is structured as follows. After reviewing related work, we will discuss the field studies on lecture notetaking. Based on these results, we will present the design of CoScribe before we finally provide results of the user study.

**Related Work**

Much research has been conducted on systems enabling users to annotate documents on *pen-sensitive displays* imitating the use of pen and paper. XLibris (Schilit et al., 1998) is a digital reading device which enables the user to create handwritten annotations with a stylus. With NotePals (Davis et al., 1999) students can take notes on a PDA during a lecture. These notes are automatically associated with the appropriate slides. Own and shared notes can be retrieved later on in a software viewer. Dynomite (Wilcox et al., 1997) combines handwritten meeting notes, made on a pen-enabled display, with audio recordings. Its so-called properties, semantic types for classifying individual handwritten annotations, inspired our tagging function. Livenotes (Kam et al., 2005) focuses on the synchronous sharing of classroom annotations on lecture slides. Each user disposes of a Tablet PC. All annotations are shared in real-time with the other members of the learning group. Hence, lecture slides form a shared workspace for several learners. Classroom Presenter (Anderson et al., 2006) and Dyknow (Berque, 2006) use Tablet PCs for communication within the classroom. Students can annotate lecture slides and submit their handwritten solutions to given questions to the instructor.

All these systems have the disadvantage that the user is limited to interacting with one single pen-sensitive display, which does not provide all affordances of paper. For this reason, several systems use *real paper* as an input medium for digital notes and annotations. Augmented paper notebooks, such as the AudioNotebook (Stifelman et al., 2001), the A-Book (Mackay et al., 2002), ButterflyNet (Yeh et al., 2006), and LiveScribe (http://www.livescribe.com), enable users to make handwritten notes on empty paper sheets and to access a digital copy of them in a document viewer. Some of these systems moreover link handwritings to additional media, such as audio recordings (AudioNotebook, LiveScribe) or photos and physical specimen (ButterflyNet). However, it is not possible to annotate already existing documents. In contrast, PADD (Guimbretière, 2003) and PaperPoint (Signer and Norrie, 2007) enable users to print a physical version of an existing PDF or PowerPoint document. Handwritten annotations made on these printouts using a digital pen are automatically added to the digital version of the document. While PADD does provide support for collaborative settings, PaperPoint supports only co-located collaboration. PaperCP (Liao et al., 2007) (an extension to the Tablet PC-based Classroom Presenter) integrates paper input with synchronous collaboration: During class, students can submit annotations they made on printed slides to the teacher or to their classmates. But each slide is printed on an own paper sheet resulting in extensive printouts. AirTransNote (Miura et al., 2007) offers functionality similar to PaperCP, but only supports taking notes on an empty page instead of annotations. While borrowing the idea of paper input, we focus on asynchronous collaboration, the visualization of shared handwritings, and mechanisms for paper-based tagging.

The PaperPDA (Heiner et al., 1999), combining a conventional paper notebook and organizer with electronic capabilities, includes interaction techniques for manipulating and tagging content on paper. It introduces the general idea of buttons which are printed on paper and enables users to create links between paper pages with adhesive stickers. PapierCraft (Liao et al., 2008) presents a set of gesture commands, like copy/paste and linking, for manipulating paper documents. PaperCP (Liao et al., 2007) expands on that and discusses other techniques for tagging paper content. These have in common that they require an additional device, like a foot pedal or a second Anoto pen. In contrast, the techniques proposed in this paper do not require this.

**Field Studies on Notetaking in University Courses**

Before designing the system, we conducted two field studies to examine notetaking and annotation processes in university courses, since notetaking in general and the use of annotation software in learning settings are not well studied (Boch and Piolat, 2005).

The first study was a quantitative questionnaire-based investigation with 408 students in four computer science (316 participants) and one pedagogy lecture (92 participants). The goal of the study was to investigate 1) how many students take notes during lectures, 2) which media are used for this purpose and why these media are preferred, and
3) the effects of taking notes on collaboration with other students. In the statistical analysis, we investigated correlations and performed $\chi^2$ tests and t-tests to identify significant group differences. All these tests were based on a level of significance of 95%. The results were as follows:

**Notetakers.** The majority of the students indicated taking notes during the lecture. However, the proportion considerably varied between the disciplines. While 93.3% of the pedagogy students took notes, only 62.3% of the computer science students did. When asked for the reasons for not taking notes, the largest group of answers to an open question considers the course slides offered by the instructor to contain sufficient information (N=16). Eight respondents indicated that taking notes distracted them from listening.

**Media use.** We further asked the students on which media they take their notes. This aspect is of central importance for the design of notetaking systems, as it should be able to support the most frequently used media.

Figure 1 shows the percentage of notetakers on single media or on combinations of several media. Both in computer science and pedagogy, traditional notetaking with a pen and paper clearly outperforms notes on a laptop. In the computer science courses, 77% of the respondents took their notes exclusively on paper. This group consists of three subgroups of roughly equal percentages which took notes either on empty sheets of paper, on printed course slides or on both of them. 8% made an exclusive use of a laptop, while 15% indicated to prefer cross-media notetaking, which combines notes on a laptop with notes on empty sheets of paper or on printed course slides. In the pedagogy course, laptop use was almost not existent. 98% took their notes exclusively on paper. The two largest groups (about 45% each) took notes either only on empty sheets of paper or combined them with printed slides or the printed course script. It is worth noting that the percentage of students taking notes on a laptop was small even though 78.6% (N=180) of the notetakers possess a laptop. Only 19.6% (N=35) of the students owning a laptop took notes on this device. The hardware equipment does therefore not seem to be relevant. Summing up, the large majority of the participants do only use paper for taking notes in lectures. The results further show that an annotation system should offer both the possibility to annotate printed handouts (e.g. slides of the lecture) and to take longer notes in free areas or on empty sheets of paper.

Taking a closer look on laptop users, we asked them about the software they took their notes with. Responses fell under two categories of almost equal frequency: Software that allows to annotate the electronic course slides (e.g. Adobe Acrobat) (N=16) or word processors and text editors (N=17). Four students indicated to use both annotation software and a word processor, three students annotated on a tablet PC, and two students employed a specific software tool for creating mind maps. These data show that for laptop users, we can identify approximately the same distribution between annotating slides and taking notes on a blank document as for the paper notetakers. The main device for the input of electronic notes is the keyboard, since only few students own a Tablet PC (1.2%).

**Advantages of media.** Besides assessing the distribution among different notetaking media, we aimed to gain information about the reasons for choosing those media as well as the advantages respondents associated with paper or electronic notes on a laptop. The results of the quantitative question are depicted in Fig. 2. They show that the
free-form flexibility was regarded as the most important advantage of notes on paper. This is followed by the fact that paper can be easily transported. As far as electronic notes are concerned, all proposed advantages were rated almost equally. Long-term archivability was rated significantly more important for laptop than for paper notes \([T=-5.935, \text{df}=234, p=.000]\). Similarly, good readability of typescript is rated significantly more important than good readability on paper \([T=-5.907, \text{df}=230, p=.000]\).

![Figure 2. Advantages of notes on paper and on laptops](image)

**Notes vs. Annotations.** The responses also indicated reasons for preferring annotations on course material or notes on blank sheets of paper. Annotations on printed course slides are regarded as advantageous, since these allow to establish a direct reference to the context by taking the notes on the place they refer to \((N=27)\). 24 respondents particularly valued that they do not need to write everything down on the slide, but only add additional important information. On the other hand, blank sheets of paper are favored because they allow to create an own structure and to note own ideas more individually \((N=12)\). Moreover, in contrast to course slides, they provide sufficient free space \((N=5)\). Three students indicated to combine notes on paper and on printed course slides in order to separate their own ideas from additional information given by the instructor.

**Subsequent use.** Respondents who took notes were asked how frequently they review and complete their notes after class and when preparing for the exam. These results allow to estimate in which phases an electronic notetaking system would be used. Mean values are depicted in Fig. 3. (Results for exam preparation relate only to the computer science courses, since in the pedagogy course, no final exam took place.) The results show that, in contrast to the wrap-up phase after class, where frequencies are rather low, students become more active when preparing for the exam.

![Figure 3. Subsequent use of notes](image)

No significant differences were found between laptop and paper notetakers. However, taking a closer look on the group which took notes on paper, our data indicate that annotations relate to more frequent follow-up activities than notes on empty sheets of paper: In review after class as well as in review and completion before the final exam, mean frequencies of respondents annotating printed course slides or the course script were .4 to .6 points higher \((p < .006)\) than of students taking notes on blank sheets.
Collaboration. We further asked the respondents to indicate their collaborative activities which make use of their notes. 54.4% (N=135) of the respondents who took notes during class indicated to use them for collaborative work. The most important point (71.1% of this group) consisted of using the notes as a basis for group work and discussion with other students. 51.1% compared their notes for completion with those of others. 45.9% gave their notes to other students or used those of others, e.g. in case of illness. Collaborative use of notes does not seem to relate to a specific medium on which respondents took their notes.

Quantity of annotations. In a second study, we analyzed annotations created by 24 postgraduate students during nine sessions of a university course. One group made handwritten annotations with pen and paper, another group typed annotations on a laptop using specific software for typewritten annotations on lecture slides. A main finding was that the use of a keyboard to type annotations seems to constrain the annotation process. Participants of the notebook group in average made considerably less annotations per participant and session (M=7.0, SD = 4.6) than those of the paper group (M=21.6, SD = 11.8). We infer that the user interface of an electronic annotation system should impose as little constraints as possible.

Semantic classification. Another outcome relevant to our system design was that many learners used semantic classification schemes to mark document passages or own annotations for later retrieval. These schemes comprise specific symbols, such as question marks, exclamation marks or asterisks. An important proportion of 11.0% of the total of 1081 handwritten annotations acted as semantic markers containing nothing but one of these symbols.

We derive the following implications for the design of CoScribe: The system shall support users in taking handwritten notes on paper and in accessing them on a computer. This combines the flexibility and the mobility of paper with the capabilities of digital support. Accounting for the highly individual annotation style, the paper-based user interface should impose as little constraints as possible on the flexible interaction with paper documents. Further, the system shall support both notes on empty sheets and annotations on printed documents. Moreover, it shall provide a mechanism for semantically classifying contents. Finally, collaboration with other learners shall be fostered.

Overview of the Design of CoScribe

We will now provide an overview of the design of CoScribe, which will be discussed in more detail in the following sections.

CoScribe enables students to 1) create handwritten annotations on printed lecture slides, on empty sheets of paper and/or on a pen-sensitive display, 2) to semantically tag annotations, 3) to bookmark document pages, and 4) to collaborate with other learners by sharing own annotations and accessing shared ones in a digital viewer. This is illustrated in Fig. 4.

![Figure 4. Overview of the design of CoScribe](image)
Our design was guided by the goal of finding an appropriate trade-off between a paper-based user interface which is as unconstrained as possible, permitting each student to maintain his or her personal annotation style, and imposing only minimal structuring, which can be leveraged for computer interpretation of informal annotations. Moreover, we judge it important not to suddenly impose major changes to the personal annotation process, which often has evolved over years. Therefore, CoScribe provides for a smooth transition from traditional paper annotations to an electronic collaborative annotation process. Moreover, the system is easily deployable in real university settings.

The system comprises a mobile part, which can be used for in-classroom annotation and only requires paper printouts of the lecture slides and a digital pen, which can be used to write normally on paper (Fig. 5). The equipment is thus highly mobile, lightweight, and comparable to that used traditionally for notetaking in lectures, so that the technology remains in the background as much as possible.

After class, CoScribe offers a combined printed and digital user interface. Students can use a PC to synchronize their annotations with the backend system, which stores them in a database on a central server. The CoScribe software viewer (Fig. 6) then provides access to a digital version of the lecture slides including own annotations and shared ones of other learners. Moreover, the annotations made by the instructor during the course (on paper printouts or on a pen-enabled display) are digitally available. At any time, the user can add or modify annotations – either on the printout using the digital pen or in the viewer using a stylus on a pen-enabled display. A print module enables the
user to print updated versions of the lecture slides including his or her own annotations, the annotations of another learner or a selection of the most relevant annotations made by several learners. CoScribe is implemented in Java and supports PowerPoint lecture slides.

**Mobile Paper-based User Interface**

**Digital Paper Technology**

For the electronic synchronization of the user actions on paper, our system relies on the Anoto technology (http://www.anoto.com) and Logitech io2 digital pens (http://www.logitech.com). During writing, the digital pen identifies both the page and its position on this page by detecting a specific, nearly invisible dot pattern with a built-in camera. An ordinary laser printer can be used to print out documents including this pattern. If used during a lecture, the pen stores its movements as vector graphics, which are synchronized with the backend system once it is docked. Alternatively, when a computer is nearby, the pen streams data continuously to the computer via a Bluetooth connection. As a consequence, the digital view can be updated in real-time.

In contrast to other solutions for capturing handwriting on paper, no manual calibration of paper sheets and no sensing device other than the digital pen are necessary. This provides for real mobile use and for naturally writing and switching between multiple paper sheets like with a traditional pen and paper.

**Flexible Annotation and Notetaking**

The printouts were designed to constrain the personal annotation style as little as possible. Handwritten annotations, including free-form content such as drawings or formulae, can be made anywhere on printouts. In addition to the printed lecture slides, the pages can optionally contain free areas (see Fig. 7 for one example printout). Our observations have shown that free areas seemed necessary to most students, because firstly the slides often do not contain enough empty space for extensive annotations and secondly it is considered important to take notes independently of the instructor's structuring.

![Figure 7. Example layout of a printed page](image)

Various printout layouts of slides and free areas are possible, providing for an adaptation to the course and the user's preferences. One physical paper sheet can contain up to four slides with or without free areas. In addition users can dynamically add empty notetaking areas when more space is needed: Additional empty paper sheets can be written on and associated with the corresponding slide by drawing a link gesture. This is a line connecting a slide printout with an empty paper sheet and has been earlier suggested in (Liao et al., 2005). All annotations made on paper are automatically transferred and displayed in the CoScribe viewer (Fig. 6).
In a typical lecture setting, the slides are annotated by the instructor and projected onto the wall of the lecture hall. For this purpose, CoScribe includes an instructor interface. He or she can make handwritten annotations on a paper printout of the slides or on a pen-sensitive display in the software viewer. These are projected in real-time and moreover can be accessed by the students in the software viewer later on. In addition, the instructor can use digital pen and paper as a remote control for the projection. By tapping with the pen on a printed slide, this slide is automatically projected.

**Paper-based Tagging using Tag Buttons**

Tagging lecture content with semantic markers is closely related to annotating: It can support individual structuring of the content (e.g. marking the beginning of new sections) and classifying passages (e.g. as important or unclear). Our field study has shown that these tagging processes occur frequently. Moreover, in a collaborative system, tagging can be used to define if an annotation should be visible to other users or not. We therefore developed two paper-based tagging mechanisms that can be easily integrated into current annotation practice.

CoScribe provides for tagging annotations with explicit semantic categories, adding a more formal representation to the informal annotations. The interaction was inspired by buttons in Graphical User Interfaces. A toolbar containing several printed “buttons” is printed in the center region of each paper sheet (see Fig. 8 and 7). Tags are associated to an individual annotation by tapping with the pen on the corresponding button(s) before or after writing the annotation. Moreover, tags can be associated or modified later on by making two consecutive pen taps on the button and on the annotation. In the CoScribe viewer the tags are visualized with specific symbols and colors (Fig. 6).

![Figure 8. Toolbar with printed tag buttons](image)

Applying tags to entire annotations instead of unstructured sets of pen strokes requires clustering these pen strokes made by one user. For this reason, CoScribe uses a clustering algorithm for handwritten input that relies on temporal and spatial information. Possible errors can be manually corrected by the users. The software viewer therefore provides two functions for splitting annotations and merging pen strokes into one annotation.

Annotations can be tagged on the two levels of semantic types and visibilities. CoScribe offers four semantic types, derived from the needs identified in our field study (important, question, to do, and correction). In addition, learners can use tag buttons to define the visibility of the annotation. An annotation can be defined to be private, visible to the members of the user’s learning group or visible to all students attending the lecture.

This slight formalization of informal handwritten annotations provides benefits both for search and retrieval of one's own annotations and for the orientation within shared annotations of other users. In addition, the semantic types act as instructional scaffolds stimulating students to perform meta-cognitive self-control processes (see for example (Scardamalia, 2001)), such as assessing where the important points of the topic are and where the student has gaps in his or her understanding and must perform further work.

This “tag button on paper” approach has several advantages over the classification means discussed in related research (Liao et al., 2007, Liao et al., 2008), which are spatial differentiation and pen gestures. Providing on the printout separate areas for each type and requiring the user to write a note in the corresponding area is impractical for annotations, which often have to be made at specific context positions on the slides. Another suggested solution consists in classifying notes by performing specific pen gestures. However, current Anoto pens cannot recognize gestures. Gesture recognition would have to be made by the backend system. Hence, it is not possible to provide feedback on success or failure of the gesture recognition in the mobile setting where pen data is not immediately transferred to the backend system. Moreover, the system would have to distinguish gestures from ordinary handwritings or drawings. For this purpose, the cited research papers suggested to use additional hardware like a foot
pedal or a second pen. In contrast, tagging with paper buttons does require only one single pen. A correct
interpretation is guaranteed, as determining the pen position on a paper button does not imply uncertainty. The pen
moreover provides graphical feedback to the user at the moment the button is tapped on. However, no graphical
feedback on the tagging is provided on the printed slide unless the user makes additional markings. Yet, the tags are
visible in re-printed versions of the slides and in the digital viewer.

Tagging with Digital Paper Bookmarks

While the “tag button on paper” approach provides for a fast classification of fine-grained objects with a set of pre-
defined classes, the second approach allows to tag entire slides with freely chosen keywords. This approach is
inspired from interaction with traditional paper and provides full feedback on paper. Adhesive page marker stickers
are a well established practice for structuring documents, marking specific passages and accessing them later on.
Impressive examples are books of law students, which often contain several dozens of bookmarks attached to the
margins of the pages. As a metaphor, bookmarks also quickly became familiar in the electronic world for marking
documents on the Web. Empirical work shows that their use positively influences the perceived ease of finding
information (Thakor et al., 2004). Moreover, bookmarks provide a means to individually structure documents and to
create structural knowledge (Jonassen et al., 1993).

We augment traditional practice by printing the Anoto pattern on adhesive stickers of different colors (Fig. 9).
Stickers can be attached to physical pages of the printed document at arbitrary positions. A simple link gesture with
the digital pen then informs the system about the position (1). Handwritten labels are electronically captured as well,
providing for free tagging with user-defined tags (2).

Once associated, the paper bookmarks are also visualized in the digital document, in the slide preview panel and in a
separate bookmark panel (Fig. 6). This latter displays a three-dimensional representation of the paper stack and
visually indicates the number of pages of the document, the current position within it and the bookmarks associated
to it (see Fig. 6 lower left). We designed this view with the goal to support cross-media navigation in settings where
the user works with both the printed and the digital version of a slide set. The view provides visual information
helping to find the corresponding paper instance of the page currently displayed in the viewer. At a single glance, the
user gets information about the approximate position of the page within the paper stack, about the nearest bookmark
and the page distance to this bookmark. She can then find the paper page by recurring to its approximate position or
by choosing an appropriate bookmark located near this page. Furthermore, the bookmark panel also supports
orientation and navigation within the digital document. Pages can be displayed by clicking on the three-dimensional
representation of the paper sheet or on a bookmark.

Visualization of Shared Annotations

A challenge with shared handwritten annotations is their clearly arranged visualization, particularly for a large
number of users. We explore two approaches:
By separating the annotations of different users into different views, each of these views in itself becomes easier to read. CoScribe provides a single-user view for each member of the user's learning group. This view can be accessed by selecting the button in the toolbar which is associated with the view of a specific user. Yet, this implies the need to manually switch between views. This becomes particularly cumbersome in larger communities. Therefore, we developed a novel visualization of collaborative handwritten annotations.

The multi-user view displays both one's own and shared annotations in an integrated manner supporting overview and access to shared annotations without the need of explicitly switching. One's own annotations are visualized as they are written on paper. However, shared annotations of other users are displayed in a condensed form to account for the restricted space within the document. Instead of the shared annotation itself, a small icon is visualized at the position of the shared annotation (Fig. 6, upper right). This icon corresponds to the type of the annotation and varies in size according to the size of the annotation. When hovering the mouse over the icon, the annotation is expanded and displayed at the correct position in its original size (Fig. 6, lower right). Shared annotations considered especially relevant can be permanently expanded (Fig. 6, annotation with grey background). An additional preview function provides for an overview on all shared annotations on a page by expanding all of them in a scaled-down manner.

Evaluation

We conducted two user studies to evaluate the appropriateness of the system both for in-classroom annotation and for the collaborative review of annotations. Our goal was first to assess if the paper interface is easy to learn and to use. Second, we explored the use of Digital Paper Bookmarks. Third, we evaluated the multi-user view for shared annotations and compared it with the single-user view.

Method

First the in-classroom setting was evaluated with 29 computer science students who used the paper interface of the system during one of their normal lectures (each lasting about 90 min.). 24 of the participants were male, five were female. Their task was to make annotations on digital paper printouts during the lecture as they usually do; in addition, we asked them to semantically classify and/or to share annotations using the paper buttons when it seemed appropriate to them.

![Figure 10. Main quantitative results of the evaluation. (All statements are transformed to their positive form.)](image)
first study. In both studies, the participants were initially trained for about three minutes on how to use the pen and the printed interface. Feedback was gathered with Likert-scale questionnaires and semi-structured interviews. To avoid positive bias, none of the students was personally known to us nor did they attend one of our courses.

Results and Discussion

Main quantitative results of the evaluation are depicted in Fig. 10. In the answers to the questionnaire, annotating with the digital pen was considered only slightly more distracting than with a normal pen, but significantly less distracting than annotating with a laptop (T=-9.0, df=26, p<.001). Although the users have had only a few minutes for familiarizing with the system and used it during one of their normal lectures, they were able to make a considerable amount of annotations. In all three lectures, a total of 1983 handwritten annotations were made. This results in an average of 68 annotations (SD=29, N=29) per user during a lecture. The participants were active throughout the entire lecture, each user annotating on average 63 % of all lecture slides.

The participants had different preferences concerning the layout of printed slides. This underscores the importance of flexible layouts. Three out of four of the participants reported in the interviews to prefer only two slides per paper sheet, as this leaves free areas for annotations. The remaining participants preferred four or more slides per page. All left-handers wished having free annotation areas to the left of the printed slides, while right-handers preferred them to the right. Positioning the free areas below instead of beside the slides was judged less appropriate, since the items on the slides are typically organized in vertical order.

The participants valued the possibility to classify annotations with semantics directly on paper. They were up to accept the additional effort of tapping on the appropriate tag buttons during the annotation process. This is reflected by a substantial percentage of 18.7 % of all 1983 annotations that was tagged with a semantic type. These findings and qualitative feedback from the interviews lend support to the assumption that paper-based classification imposes little overhead. The two most frequently chosen types were ‘important’ (12.5 %) and ‘question’ (3.2 %). Although tagging with visibilities was performed only for a small percentage of annotations (2.4% private and 1.6% public; the default was group visibility), the interviews indicate that this functionality is important. Several participants expressed the wish to define individual types like ‘to revise’, ‘definition’ or ‘example’.

The bookmark-related quantitative responses show that users judged bookmarks being helpful for finding specific pages both in their paper version and in the software viewer. Several participants particularly valued the possibility to fade away from the given structure and instead to create an individual structure of the document.

In the questionnaires, the participants expressed a clear preference for the multi-user view over several single-user views when seeking a quick overview of all comments. This result is highly significant (T=-6.64, df=10, p<.001). They also judged this view slightly better for quickly finding specific shared comments. The multi-user view remained clearly arranged with four learning group users in the evaluation setting. We assume this view to scale to a larger user community within university lectures as well, since, according to our first experiences, the total number of users does hardly affect the number of shared annotations a specific user can retrieve. This relates to the fact that by default, annotations are shared at the intermediate learning group level and that only a very small fraction of annotations is published to the entire community.

Conclusions

Based on the findings of two preliminary field studies, we have developed CoScribe, a concept and system for the collaborative annotation of lecture slides on digital paper. CoScribe significantly reduces the gap between the use of paper, which is frequent in university lectures, and electronic annotation systems. The results of a user evaluation indicate that the combination of paper with computing is very promising for supporting student annotation and collaboration.

Although the system was designed with a special focus on university lectures and evaluated in lectures where the students were physically present, we envision it to be helpful in other settings as well, encompassing distance learning and the collaborative creation and editing of presentation slides. CoScribe provides strong support for
autonomous and group-based learning. More enhanced forms of tagging and linking of documents are under
development and will be integrated in a future version of CoScribe

Now that digital paper can be used as an input medium for collaboratively creating user generated content in large
communities, we envisage studying how principles of the Web 2.0 can be applied to paper-based computing in order
to foster learning processes. For example, by performing automatic analyses and statistical aggregations on paper-
based contents, particularly relevant annotations of other users could be highlighted. Digital Paper Bookmarks could
be a basis for combining electronic social bookmarking with physical documents. This could lead us a further step
towards bridging the paper world with state-of-the-art collaborative technologies.

Acknowledgments

We thank our students Simon Olberding and Roman Lisserman for their help with the implementation of the
prototypes. We are grateful to Michael Deneke and Oliver Glindemann for their advice and support. Moreover, we
thank the reviewers of ICALT 08 for their helpful comments on an earlier version of this article. This work was
funded by the German Research Foundation as part of the Graduate School on Feedback-Based Quality Management
in eLearning (DFG-GK-1223).

References

presenter - a classroom interaction system for active and collaborative learning. In Workshop on the Impact of Pen-based
Technology on Education (WIPTE), April 6-7, 2006, West Lafayette, IN, USA.

Berque, D. (2006). An evaluation of a broad deployment of dyknow software to support note taking and interaction using pen-


& Schilit, B. N. (1999). Notepals: lightweight note sharing by the group, for the group. Proceedings of the SIGCHI conference on


Acquiring Structural Knowledge, Hillsdale: Lawrence Erlbaum.

and augmented note-taking in lectures. Proceedings of the SIGCHI conference on Human factors in computing systems, New

(2), 147–172.


Computer Society, 175-179.


A Methodological Approach to Encourage the Service-Oriented Learning Systems Development

David Díez, Alessio Malizia, Ignacio Aedo, Paloma Díaz, Camino Fernández and Juan-Manuel Dodero

1DEI Laboratory, Computer Science Department, Escuela Politécnica Superior, Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés, Madrid, Spain // Tel: +34 91 624 99 60 // Fax: +34 91 624 91 29 // david.diez@uc3m.es // alessio.malizia@uc3m.es // ignacio.aedo@uc3m.es // npaloma.diaz@uc3m.es
2Mechanic Engineer, Computer Science and Aeronautic Engineer Department, Universidad de León, Campus de Vegazana s/n, 24071 León, Spain // Tel: +34 987 29 30 64 // Fax: +34 987 29 30 65 // camino.fernandez@unileon.es
3Lenguajes y Sistemas Informáticos, Escuela Superior de Ingeniería, Universidad de Cádiz, c/ Chile 1, 11002 Cádiz, Spain // Tel: +34 956 01 57 84 // Fax: +34 956 01 51 39 // juanma.dodero@uca.es

ABSTRACT

The basic idea of service-oriented learning is that a learning environment should be conceived as a set of independent units of learning packaged as learning services. The design, development and deployment of a learning system based on integrating different learning services needs both a technological platform to support the system as well as a methodology that manages the development process. Such technological platform is provided by the Service-Oriented Learning Architecture (SOLA). SOLA proposes an architecture style, which defines how to deploy several units of packaged-learning services. The purpose of this paper is to review the SOLA concept and to present a specific method based on specifying instructional domains, to complement the SOLA approach. Our method suggests the definition of a feature model, which represents the characteristics of the instructional domain, in order to address the identification and classification of candidate services. The approach here proposed facilitates the definition of candidate-learning services in order to enhance the integration of such services.

Keywords

Learning service, computer-based learning, service-oriented learning, instructional domain, feature modelling

Introduction

The definition of standards and specifications in eLearning aims at improving reusability, portability and interoperability (Friesen, 2005). These qualities should apply both to the learning system and to the learning components used by the system. So far, the standardization process has been primarily focused on defining standards that specify contents and their structure. This strategy is suitable to achieve reusability and portability, but it has shown limited capacity to support interoperability; especially, dynamic semantic interoperability. Dynamic semantic interoperability includes the capability of learning systems and learning components to exchange information and to use the information that has been exchanged without any additional user interaction.

The solution to this weakness might lie in service-oriented learning. In this approach, a learning system can be defined as a set of units of learning packaged as services. In order to accomplish interoperability, each learning service is assigned its own functional context (learning objectives) and it is encompassed by a set of capabilities (inputs, outputs and preconditions) related to its educational context. An eLearning platform based on services will be able to dynamically discover and assemble learning components to achieve the desired specific learning purpose (Dagger et al., 2007). The assembling of services implies the coordination and orchestration of services (Erl, 2005). Service coordination and orchestration enable services to be strung together following predefined patterns. These patterns establish how services interoperate to achieve a set of learning objectives. Hence, the development of a learning system based on integrating services needs an analysis process that identifies such patterns in order to know which units of learning should be encapsulated by each learning service. Despite its relevance, there are no current suitable proposals that guide the processes of definition and composition of services.

In this paper, we postulate a top-down analysis phase based on domain analysis (Prieto-Diaz, 1990) and feature models (Czarnecki, 1998) in order to address the process of identification and classification of learning services. The proposed method is an analysis approach conceived to improve the development of computer-based learning. Such method applies the principles of domain analysis to the field of automating learning in order to systematize the development process of computer-supported learning systems. In particular, within the service-oriented learning
context, our approach proposes a set of mechanisms that, based on an instructional domain model, address the identification and classification of candidate services. Later, at the design phase, such candidate services will be organized into specific learning services.

The rest of the paper is organized as follows. Section 2 reviews the concept of service-orientation and its usefulness in the development of computer-supported learning systems. Section 3 introduces an analysis methodology conceived to support service-oriented learning. Finally, conclusions and a set of recommendations for further work are drawn in the last section.

Service-Oriented Learning

A learning system may be conceived as a set of instructional activities, contents and learning resources, structured in a way that a number of learning goals can be achieved. Within the computer-supported learning context, this set of learning components can be managed as units of learning. Each unit of learning is in turn as a service. The specific application of this concept to the development of computer-supported learning systems constitutes the key principle of the service-oriented learning paradigm. The basic idea underlying service-oriented learning is that a learning environment should be understood as a set of independent units of learning, packaged as learning services.

The main contribution of service-oriented learning is oriented to the reusability of learning components and the development of flexible eLearning platforms (Westerkamp, 2006). However, this activity is not trivial as it requires the development of concrete mechanisms that must guide the development of the learning system and manage its later deployment. Subsequent sections delve into the service-oriented learning approach. Emphasis is put into the definition of the ‘learning-service’ concept, the review of the notions of Service-Oriented Learning Architecture (SOLA) and the introduction of one of the most important phases of the development process: the analysis phase.

The concept of learning service

A learning service is usually understood as a supporting activity (Wilson et al., 2004). Authentication, logging or authorisations are examples of learning services. Within the service-oriented paradigm, a service relates to system functionality: a service is created to resolve a set of functional goals. Since an instructional process is designed to achieve some educational objectives, we will use here the term learning service in the sense of service-orientation and, therefore, we propose the following original definition:

Learning service. It is a set of learning activities, learning resources and support facilities that are used to satisfy some specific learning objectives.

A learning service is composed by one or many operations. Each operation represents either a learning activity (e.g. question and answer activity) or a support facility (e.g. e-mail). As far as a learning object is concerned, a learning object is conceived as any reusable digital resource that is encapsulated into either a learning service or a message among services. In this approach, there is a relationship between learning services and units of learning (UoL). Both of them encapsulate a unit of business logic, a delimited piece of education. However, they differ in their autonomy. The logic governed by a service resides within an explicit boundary. The service has control within this boundary and is not dependent on other services or tools for executing its functionality. A learning service can be used by several systems, being each of these systems potentially deployed in different contexts. However, a learning service can be executed successfully within the different systems independently of the context of use. On the other hand, a UoL (IMS-LD, 2003) does not have control over whatever learning content it represents; thus, a UoL is restricted to content modelling with the purpose of facilitating its later reuse. Hence, UoL needs specific management utilities provided with the runtime system (Olivier & Tattersall, 2005). For instance, before executing a UoL it is necessary to map out all the context-specific resources identified at the design stage. The way in which resources are mapped is determined by the runtime system. This restriction reduces the autonomy and interoperability of UoLs.

Service-Oriented Learning Architecture

The key feature of a service is to provide a specific behaviour that can be applied by different technological platforms (Wiegers, 2003). In order to fulfil this requirement, a baseline architecture technology is needed. The Service-Oriented Architecture (SOA) represents an example of such kind of artifact. SOA is a term that represents an
architecture style in which automation logic is decomposed into distinct units of logic. Applying this approach to the educational context, Service-Oriented Learning Architecture (SOLA) has emerged. SOLA is defined as “an IT strategy that organizes the discrete functions contained in an educational network into interoperable, standards-based services that can be combined and reused to meet teaching, learning, research and administration needs” (Clark & Booth, 2006).

SOLA provides a technological platform to integrate independent learning services. One of the most interesting aspects of SOLA is the ‘layer of services’ concept. A layer of services abstracts a specific operational aspect of a system (Erl, 2005). In particular, three layers of services are identified:

- **Application service layer.** Services that reside within this layer can be referred to as application components and operations. Authentication or authorisation can be included within this layer.
- **Business service layer.** While application services are responsible for representing technology and application logic, the business service layer is solely concerned with representing business logic. Units of learning and learning object packaged as services should be integrated in this layer.
- **Orchestration service layer.** The orchestration service layer consists of one or more services that compose business and application services taking as a reference an instructional process. Course management, sequencing or learning flows can be included within this layer.

According to the SOLA paradigm, the development of computer-supported leaning systems is based on identifying and composing different learning services within the appropriate layer. However, this task can be difficult, particularly if developers are not provided with mechanisms to support the identification and the composition process integrated within the educational requirements of the system. In this context, we propose a framework oriented towards supporting an analysis process that will assist in the definition of a business model known as the instructional domain model. The instructional domain model represents the domain knowledge structures that can be used to classify the learning services as described in the next section.

**A Service-Oriented Learning Analysis**

A learning system is the product of an instructional engineering process (Paquette, 2004). The life-cycle of the earlier instructional engineering methods resembles the phases of the traditional software engineering process: analysis, design, development and evaluation. Following the opinion of diverse authors (Alessi & Trollip, 2001; Preece et al., 2002; Wiegers, 2003), the analysis phase plays a preponderant role. Hence, we will review the analysis phase.

Service-oriented analysis focuses on identifying and modelling learning services. Service-oriented analysis essentially identifies service capabilities, which in turn are grouped into logical groups. Such logical groups are later assembled into service candidates. This process can therefore be guided by two main analysis strategies (Erl, 2005): top-down and bottom-up strategies. The top-down strategy promotes the formal definition of business models prior to modelling service boundaries. The bottom-up strategy is based on the delivery of application services on an ‘as needed’ basis. In this paper we postulate a top-down analysis phase.

The selection of a top-down analysis strategy has been motivated by the following issues:

- The field of computer-based learning may be considered as an isolated and stable context. It means that the learning systems of the context are guided by theories and well-known principles.
- It is possible to identify reliable sources of information that establish the domain knowledge.
- The development of computer-supported learning systems requires the participation of both system analysts and domain experts - instructional designers, contents experts and teachers -.

The combination of these issues recommends the definition of a generic business model that integrates current knowledge on instructional domain. This generic model will be used as a reference by the analysts during the specification of a new system. The use of a generic business model facilitates the recognition of common patterns, being these patterns the basis to identify and classify candidate services.
The Methodological Approach

Instructional Domain Analysis is a top-down methodology based on applying the domain analysis paradigm to the instructional design. Such a methodology integrates knowledge from both software engineering and instructional design in order to propose a systematic analysis process of computer-supported learning systems. Our approach provides two software artifacts in order to systematize the development of a computer-supported learning system: a domain model and an analysis process based on such a model. On the one hand, the method allows arranging a specific model that represents the knowledge of an instructional domain. This domain model facilitates the communication among participants in the development process and promotes the reusability. On the other hand, the method provides an analysis process based on the domain model to study a computer-supported learning system. A domain model defines the instructional domain semantic; therefore, such a model can address the system analysis process, improving its completeness and avoiding its inconsistency. Moreover, and related to service-oriented learning, our approach proposes the definition of a feature model, which represents the characteristics of the instructional domain so as to address the identification of candidate learning services. A candidate learning service can be defined according to its characteristics; hence, feature modelling is an appropriate artifact to identify a candidate service. Prior to explaining the proposed analysis process, a set of concepts about instructional domain analysis will be presented.

Instructional Domain Analysis

Instructional Domain Analysis is an analysis approach based on defining an instructional domain model that specifies the domain knowledge. There are different artifacts to represent the domain knowledge like ontologies or domain-specific languages. The method here presented proposes feature modelling as a domain knowledge representation technique. Feature models are particularly useful artifacts to improve reusability. The reason is that reusable components contain inherently more variability that specific components, and feature modelling is the key technique for identifying and capturing such variability (Czarnecki, 1998). The following paragraphs present different concepts related to the application of the principles of domain analysis to instructional design.

Domain analysis. Domain engineering is a common activity within software engineering conceived to ease reusability (Clements & Northrop, 2001). Domain engineering proposes the systematic creation of domain models to guide development. A domain model is defined as the set of elements or properties of a system within a particular context (Prieto-Diaz, 1990). As shown in figure 1, the development process proposed by domain engineering establishes three phases, being domain analysis the first of them.

![Figure 1. The development process based on domain models (Czarnecki, 1998)](image)

Domain analysis aims at studying a specific application domain in order to define a model that represents the knowledge of that domain. This working methodology allows a priori-scoping of the problem by identifying those
elements or areas of interest that determine the development process of domain systems. The ultimate purpose of domain analysis is to treat domain systems as a whole that permits the reuse of components and experiences. Prior to applying domain analysis to computer-based learning, it is necessary to establish a definition of instructive domain and to determine the information sources required to establish domain knowledge.

**Instructional domain.** The first aspect to be defined concerns the domain and its scope. In particular, we define an instructional domain as:

*Instructional domain. The set of educational scenarios that share contents, an area of work, as well as pedagogical objectives and a specific instructional form. The magnitude of the educational contents considered must make sense by themselves.*

The scope of a domain is determined by the set of systems that satisfy a common functionality. In the case of instructional systems, the functionality to be solved is understood as the systematization of the learning process (Tennyson, 1995). Given that the learning process depends on the instructional form (which is in turn conditioned by the contents, the context and the pedagogical objectives), these factors have been taken as essential for the demarcation of a domain. As for scope issues, the domain considered must go beyond the domain of a certain course or subject and it must be more reduced than the instructional model or the group of subjects taught within a particular field of knowledge. An example of an instructional domain is the learning and teaching programming domain. Such a domain compiles a number of educational environments, allowing the instruction in computer programming through the resolution of problems and situations that imitate the real world.

The study of domains requires reliable information sources that compile the domain knowledge. In the case of instructive domains, the information sources available are:

- **Instructional design theories.** Instructional design theories offer formalized knowledge about most suitable learning methods according to the type of learning or learning conditions.
- **Learning patterns.** They show examples of a learning method in course or used within a specific context. They can be considered as instructional models.
- **Domain experts.** Instructional designers, experts in contents, and teachers with long experience in instructional domains.

These information sources complement with learning systems and instructional resources related to the contents of the domain at stake. The combination of all these information sources allows the study of instructional domains and the representation of their knowledge through domain models. In case of our approach, the definition of a domain model is based upon feature modeling.

**Feature modelling.** Feature modelling captures the general facilities of an application in a domain in terms of features. In our work, a feature is defined as:

*Feature. A distinguishable characteristic of a system that is relevant to the instructional process in a specific instructional domain.*

The system characteristics can be categorized according to their usefulness in the development process. Taking into account the features categories defined by the FORM method (Kang et al., 1998), we have classified the instructional features into four categories:

- **Capability.** Features that are associated with the operations or functions provided by the system. It corresponds to contents, learning activities and units of learning. Taking as an example the learning and teaching programming domain, as shown in figure 2, contents would be divided into ‘programming strategies’ and ‘programming knowledge’. Examples of programming knowledge are ‘data types’ and ‘control flow statements’.
- **Operation environment.** An operation environment feature represents attributes of the environment in which a system is used. It comprises support activities, platforms, tools and course management systems used in the domain. For instance, a support activity can be classified as ‘authoring’, ‘monitoring’ or ‘feedback’ and, within these, ‘surveys elaboration’, ‘tasks mining’ and ‘results publication’, respectively.
- **Domain technology.** This category represents instructional design details. In an educational context, domain technology features compile the characteristics of potential students, instructional theories and the most suitable learning methods for the domain. In case of the learning and teaching programming domain, students would be classified as ‘beginner’, ‘qualified’ or ‘advanced’; regarding the learning methods, features as ‘visualization
learning’ or ‘discovery learning’ could be identified. The ‘discovery learning’ feature is realized for other features as ‘problem-based learning’, ‘project-based learning’ or ‘game-based learning’.

- Implementation technique. This category represents implementation details used to deploy a learning system. It compiles the instructional standards, specifications and support services used in the domain. Examples of features of this category would be ‘authentication services’ or ‘messaging services’; similarly, a messaging services can be defined as ‘synchronous’ and ‘asynchronous’.

![Figure 2. Domain-feature model](image)

At this point, the difference between support service and learning service should be explained. A support service represents an isolated and independent action (Wilson et al., 2004) that can be catalogued into the implementation technique category; however, according to the definition previously provided, a learning service embodies a set of activities and resources in order to achieve a specific instructional goal. A learning service is a multifaceted concept that compiles a number of features spreading amongs different perspectives, so a learning service cannot be easily classified into a specific category. With the purpose of classifying a learning candidate service, our approach proposes a relationship between the problem space (feature model) and the solution space (layers of services in the SOLA architecture). In particular, we can identify relationships between the implementation technique category and the application layer; the capability features and the business layer; and the operation environment features and the orchestration service layer. The domain technology refers to the instructional process; hence, the domain technology category does not describe learning services properties but the business process that manages the learning services of the system. This relationship allows the classification of candidate learning services according to the type of features defining them - a candidate learning service should be classified into the service layer associated with the category of its preference features.

**The Analysis Process**

The analysis process defines a set of activities in order to systematize the specification of a computer-supported learning system. Our approach aims at identifying and modelling candidate services taking as reference a feature model. Thus, as shown in figure 3, the analysis process is broken down into two stages: *domain analysis* and *service analysis*. The first stage involves the collection of the knowledge that is to be represented in a feature model; the
knowledge is provided by diverse information sources. The second stage uses the feature model to identify, classify and compile the candidate learning services.

**Domain Analysis.** The purpose of the domain analysis is to collect relevant domain information and to integrate such information into a domain model. The final goal is to specify domain learning commonalities so as to identify learning services. In order to capture the general facilities of a domain and facilitate the identification of candidate learning services, our approach is based on feature modelling. As shown in figure 4, the domain analysis is an evolutionary process divided into three activities: domain scoping, domain identifying and domain modelling. The domain model is dynamic, that is, it can be refined continuously over its life-cycle by adding new domain knowledge. As more knowledge is added to the model, it becomes more precise and useful. With the purpose of achieving this objective, the domain analysis process should be repeated until the domain model is validated.

![Figure 3. An analysis process based on domain engineering principles](image-url)

![Figure 4. Domain analysis process](image-url)
Domain scoping. A domain is a logic abstraction that groups particular set of courses or areas of knowledge. Thus, the choice of a domain should be driven by the contents and pedagogical objectives. Taking into account these ideas, domain scoping compiles the following tasks: (1) Determining the discipline. The discipline is determined through the review of contents. The use of bibliographic catalogues and the syllabus of courses is recommended; (2) Review of the type of instruction to be developed. This task allows the establishment of the educational paradigm of interest; (3) Defining the domain. The combination of the discipline and the educational paradigm establishes a number of specific educational scenarios. The sum of these scenarios makes up the instructional domain of interest; and, (4) Validation of the domain. Both experts and instructional designers are useful to determine the validity and coherence of the domain. Taking into consideration the steps mentioned above, the Computing Classification System catalogue can be used to determine the computer programming discipline. Once the discipline has been identified, some information about the discipline will be reviewed in order to recognize the most common way of teaching in that discipline. According to different reviews and discussions (Milne & Rowe, 2002; Robins et al, 2003), computer programming should be taught following the principles of constructivist paradigm (Jonassen, 1999). Finally, the domain scope will be validated by domain experts.

Domain identifying. Once the domain has been scoped, the features of the domain should be identified. Feature identification is guided by ‘feature starter sets’ (Czarnecki, 1998). Feature starter sets define the collection of features suitable for the domain. In our method, such sets of features correspond to the four categories previously enumerated. For each category, compiled domain knowledge is reviewed, differences and similarities between domain entities are recorded, and specific features are identified. The activity is divided into three tasks: (1) Defining the domain vocabulary. Terminology from both the instructional domain and the computer-supported learning context is compiled. This task is focused on compiling the contents of the discipline of the domain, the instructional activities and methods associated with the instructional paradigm, and the common educational objectives of the instructional domain; (2) Identifying the domain concepts. Based on the terminology defined, domain concepts are identified. The concepts are related to the participants of the instructional process, contents, instructional technologies, etc.; and, (3) Identifying features. For each of the domain concepts, properties are identified and grouped into the features types previously described. First of all, the domain concepts (participants, contents, educational objectives, activities, etc.) are classified into the appropriate category. Secondly, based on these concepts, the most significant properties of the domain are identified as features. A feature can be described by other more specific or detailed features. Finally, relationships between features are identified and documented.

Domain modelling. The features of the domain are integrated into a feature model. A feature model organizes features into a hierarchical diagram that allows classifying such features as mandatory, alternative and optional (Kang et al., 1990). The feature model is composed of several feature diagrams. Each of these diagrams has a tree structure that uses a concept as root and the features of such concept as branches. For instance, as was shown in figure 2, the contents are related to system functionality; hence, the contents are classified into the ‘capability’ category as a root. Then, the different contents of the domain and their properties are represented as children of the ‘contents’ root. In addition to the feature model, a set of business restrictions rules amongst features should be defined. Such rules define the semantic constraints existing between features that are not gathered in the feature diagram. For instance, we can define a relationship between the instructional method and capability features: problem-based learning is advisable so as to acquire knowledge on code implementation. The formulation of restriction rules is based upon the features involved and the type of restriction between those features. i.e One feature requires another feature, one feature excludes another feature, or one feature suggests another feature.

Service Analysis. Service analysis is the process of identifying and classifying candidate learning services. A candidate service corresponds to a logical context. Such logical context is defined by a set of one or more semantic-associated features. As shown in figure 5, this stage is composed of the following steps: system scoping, service identifying and service classifying.

System scoping. The first activity aims at defining the scope of the required computer-supported learning system. The system scope is defined by educing the instructional requirements and the operational conditions of the system. This set of requirements and conditions are used, in concert with the domain model, by the following activities of the process to identify and classify the candidate services.
Service identifying. This activity addresses the identification of candidate services based on the result of the domain modelling. The first task is to select a set of features of the domain model that describes the learning system under development. For the selection of features some guidelines can be considered: (1) First of all, an effective method for identifying such a feature set is based on the four categories previously enumerated, i.e. first considering domain technologies, then capabilities, and finally the operation environment and the implementation techniques; (2) Secondly, restriction rules set up constraints and optimal selections, i.e. if two features may not be held concurrently the selection of one of them rules out the other. Once the system configuration is defined, the semantic of features is reviewed in order to recognize logical contexts. The identification of a logical context is addressed by the educational objectives of the system. Based on these educational objectives, the domain concepts related to learning activities and instructional resources are reviewed. The outcome is a logical combination of learning components that can be defined as a candidate service. For instance, learning the syntax of a programming language is an educational objective, which requires the resolution of programming exercises. If the assessment of such exercises is based on a continuous evaluation method, a portfolio will be needed. In this context, portfolio assessment and programming exercises are functionally related; thus, an ePortfolio that supports the management of programming exercises may be considered as a candidate service.

Service classifying. The classification of candidate services is guided by the feature categories previously enumerated. A candidate service can be regarded as a logical group of features; however, all features of the group do not necessarily have the same priority or responsibility. A candidate service should be classified into the service layer associated with its preference features. For instance, an ePortfolio service should be classified into the application layer because its main feature (assessment procedure) belongs to the implementation technique category. The result of the service analysis process is a classification that represents the candidate learning services grouped by layers. Such classification will be used as a starting point to specify the service contracts required. Thus, for instance, and taking as a reference the feature model depicted in figure 2, it is possible to identify the need for message and validation services that allow the achievement of peer tutoring activity. The first two services, because of their correspondence to the features of the implementation technique category, should be classified into the application layer; in the same fashion, the peer tutoring activity should be designed as a service embedded in the business layer.

Conclusions and further work

The application of service-orientation to instructional design increases the interoperability and flexibility of computer-supported learning systems (Dagger et al., 2007). According to this idea, an approach for the analysis of service-oriented learning has been presented in this paper. This approach proposes a systematic process to discover learning services taking as a reference an instructional domain model. By examining a family of related domain knowledge, it is possible to obtain a set of reference models expressed in terms of ‘features’. The model that captures the commonalities and differences is called ‘feature model’, and it is used to support both the engineering of learning services and the specification of a new computer-supported learning system. Furthermore, the use of a method based
on domain analysis establishes the baseline for the definition of a complete service-oriented development method. In this case, the existence of groups of domain features would be related to templates of services. The approach here presented has been evaluated in three ways – (1) an analytic evaluation of the features of the approach; (2) a proof of concept; and, (3) expert opinion. The results obtained with all these evaluation methods are satisfactory as a result of which the approach can be considered correct, useful and effective for the analysis of computer-supported learning systems. In particular, best results have been obtained when the approach has been used to analyse the instructional domain and to encourage the reuse of components. Nevertheless, and with the purpose of guaranteeing the validity of the approach, it will be necessary to deploy the method in diverse educational scenarios and to check its capability to identify learning services.

Finally, future work will be directed to facilitating the use of the method. With this purpose in mind, a graphic tool capable to support the analysis process, and more specifically the activities involved in domain modelling, and the identification and classification of candidate services, will be developed. In the same vein, and in order to improve the method, testing activities of the domain model will be included. The inclusion of these activities will allow the validation of domain models. These testing activities must engage two tasks – the verification of artifacts and the validation of the content of those artifacts. Expert opinion will play a fundamental role in instructional domains.

Acknowledgements

This work is supported by the Moduweb project (TIN2006-09678) funded by the Spanish Ministry of Science and Education (MEC).

References


Competence-related Metadata for Educational Resources that Support Lifelong Competence Development Programmes

Demetrios G. Sampson
Department of Digital Systems, University of Piraeus, Greece & Advanced e-Services for the Knowledge, Society Research Unit, Informatics and Telematics Institute, Centre for Research and Technology Hellas, Greece // sampson@unipi.gr // sampson@iti.gr

ABSTRACT
In the context of the emerging paradigm of Lifelong Learning, competence-based learning is gradually attracting the attention of the Technology-Enhanced Learning community, since it appears to meet the 21st Century learning and training expectations of both individuals and organisations. On the other hand, the paradigm of Learning Objects - as a way to organize, manage, offer and use digital learning resources - retains, to a large extent, its initial anticipations and, supported by Educational Metadata, it is still influencing Technology-Enhanced Learning system design. However, there not exist yet, a commonly accepted proposal for metadata that could describe competence-related characteristics of educational resources that support Lifelong Competence Development Programmes. In this paper we target addressing this issue, that is, we identify and study the main issues related to the competence-relevant characteristics of educational resources, taking into critical consideration the current state-of-the-art information models for competence descriptions (such as IEEE RCD and HR-XML) and we propose an IEEE LOM Competence-based Application Profile that can be used for tagging educational resources in a competence-meaningful manner.

Keywords
Competence Description, Learning Objects, Educational Metadata

Introduction
In the context of the emerging paradigm of Lifelong Learning, competence-based education and training is gradually attracting the attention of the Technology-Enhanced Learning community, since it provides important benefits for both individuals and organisations. At the individual’s level, a competence-based learning approach may help in identifying and targeting competences that need to be developed in order for an individual to reach certain levels of competences defined by an individual career plan and/or by the human resources department of an organization for a given job description. At the organizations’ level, competence-based training bares the potential for designing competence development programmes that targets organisational performance improvement and enhances human resource potential (Hustad, Munkvold & Moll, 2004).

A typical Competence Development lifecycle aims at the continuous enhancement and development of individual and/or organizational competences and consists of the following key steps: (a) the creation of a reference competence description through the identification of required job and task roles together with their associated expected competences and competence levels, (b) the assessment of existing competences at individual or/and organisational level, (c) the gap analysis between existing competences and the required competences for a specific job or task role, (d) the definition of competence development programmes or units of programmes to minimize the identified gaps and (e) the continuous performance monitoring and assessment to confirm improvement (Sinott et al., 2002).

Competence Development Programmes refers not only to Training Programmes that lead to some kind of formal recognition (that is, certificates or degrees), but also to informal learning activities which facilitate competences’ acquisition by practice rather than intentional learning (European Commission, 2001; Dodero et al, 2007). Learning Activities are defined as “the explicitly designed or loosely performed activities that are directed at the attainment of an explicit or implicit learning objective” (Koper & Specht, 2007; Prins et al, 2008) and they are supported by appropriate educational resources (typically referred to as learning objects). As a result, the issue of finding, selecting and assembling digital resources which are suitable for the particular learning activities in hand, is a key issue in the learning technologies literature (Ullrich, 2008).
In this context, Learning Objects can be defined as “any digital resource that can be reused to support learning” (Wiley, 2000) and they are typically described with metadata using the IEEE Learning Object Metadata (LOM) (IEEE LOM, 2002), that is, an IEEE standard for the explicit description of educational resources. The IEEE LOM provides a hierarchy of properties for learning resources in nine different categories which include technical, educational and other characteristics. However, the descriptive nature of the IEEE LOM metadata has lead to problems with lack of precision in description within the given metadata categories (Sanchez-Alonso & Froesch-Wilke, 2005) which does not facilitate efficient machine-based search and selection of learning resources. Furthermore, IEEE LOM does not include metadata elements that are specifically designed for the description of learning resources in terms of their relevance to Competence Development Programmes. In particular, it does not provide metadata elements that would describe the suitability of a given learning resource for supporting learning activities that are designed to meet competence-based educational objectives within competence development programmes.

This issue has been identified as a potential problem and there are some works that aim to address it, that is, to propose extensions of the IEEE LOM in order to support the description of competence related characteristics of learning resources (Sanchez-Alonso & Sicilia, 2005; Van Assche, 2007; Ng & Hatala, 2007). More specifically, Sanchez-Alonso & Sicilia (2005) proposed an normative metadata description of learning objects aiming to describe their intended use in a given learning context (which includes the description of a specific learner profile) so as to facilitate the acquisition of a certain competence at a given level (Sanchez-Alonso & Sicilia, 2005). Frans Van Assche (2007) proposed to describe learning objects based on their suitability for supporting competence-based educational objectives of national/regional curricula and he described two ontologies for this purpose, one that is based on a revised Bloom’s taxonomy of targeted objectives and one that aims to describe specific subject domains (Van Assche, 2007). Ng & Hatala (2007) proposed an ontology-based modeling of competences representing their definition, knowledge reference, evidence and level of proficiency, and they propose to use it for tagging learning resources in regard to their competence-based objectives. These proposals typically exploit the IEEE LOM element [9.1 Classification.Purpose], however, they do not take into consideration the current state-of-the-art in competence description modeling through the specifications proposed by international working groups such as IMS (IMS RDCEO, 2002), IEEE (IEEE RCD, 2004) and HR-XML (HR-XML, 2006).

In this paper we target addressing this problem, that is, to identify and study the main issues related to the competence-relevant characteristics of learning resources taking into consideration a critical view on competence modeling as expressed in the current state-of-the-art information models for competence descriptions, and propose an IEEE LOM Competence-based Application Profile that can be used for tagging educational resources in a competence-meaningful manner. The paper is organized as follows. Following this introduction, Section 2 presents the concept of competence, identifies the key dimensions of competence, and studies the current initiatives on modeling competencies. Section 3 describes the proposed approach for defining an application profile of the IEEE LOM standard (IEEE LOM, 2002) and Section 4 presents the proposed competence-based application profile of IEEE LOM standard produced as a result of the above described process. Finally, we present a full example from a real life case study (namely, the e-Access2Learn Project) to demonstrate and validate our proposal, and discuss our conclusions and ideas for future work in this field.

Theoretical background: Competence Definitions and Models for Describing Competences

Today, competences are proved to be a critical tool in human resource management, vocational training and performance management. However, despite the fact that competences are an important tool for various fields of application, the research community has not agreed to a commonly accepted definition of the term resulting to multiple interpretations (Boon & van der Klink, 2002; Delamare & Winterton, 2005; Sampson & Fytros, 2008a). The competence concept was originally developed in Psychology referring to the individual’s ability to respond to certain demands placed on them by their environment. Sampson & Fytros (2008) provided a review of the main definitions of competence in literature of different application fields, in an effort to provide a thorough understanding of the different aspects that this term involves, and they presented examples of competence definitions in the field of human resource management and in the field of vocational training and education (Sampson & Fytros, 2008a). Based on the analysis of the different competence definitions presented, three core dimensions of the term “competence” were identified, namely, (a) the individual’s characteristics, which refer to a set of characteristics such as knowledge, skills, attitudes, abilities, behaviors, traits, values, motives, self-concepts, aspects of one’s self-image, social role and/or self-control; (b) the individual’s competence proficiency level, which are used to classify
competences at specific levels, according to the performance of the individual when demonstrating the competence by an action; (c) the context in which the individual’s competence is applied, which may refer to a specific area of a job, to an occupation or function, to a life outcome, to work-related situations, to a specific situation, or to a specific task (Sampson & Fytros, 2008a). In this paper, we adopt the generic definition of the term “competence” given in (Sampson & Fytros, 2008a): a competence can be defined as a set of personal characteristics (e.g. skills, knowledge, attitudes) that an individual possess or needs to acquire, in order to perform an activity within a specific context, whereas performance may range from the basic level of proficiency to the highest levels of excellence.

Competencies need to be formally modeled so as to be able to interchange competencies description between systems. To this end, international specifications for competence description, such as the IMS RDCEO (Reusable Definition of Competency or Educational Objective) (IMS RDCEO, 2002), the IEEE RCD (Reusable Competency Definitions) (IEEE RCD, 2004) and the HR-XML Competencies (Measurable Characteristics) (HR-XML, 2006), have been recently proposed.

Mapping the elements of IMS RDCEO to those of the HR-XML specifications indicates that both specifications provide: (a) identification of the competence, (b) title of the competence, (c) description of the competence, (d) definition of the competence, (e) taxonomy of the competence, (f) personal information, while HR-XML adds elements for measurable evidence and measurable weights and importance levels (Sampson, Karampiperis & Fytros, 2007).

A careful examination of these specifications reveals that they do not included in their scope important dimensions of the generic competence model. Thus, based on the key dimensions of competence, Sampson, Karampiperis & Fytros (2007) have identified the following issues:

a) The notion of competency itself is not detailed. However, competence modeling should anticipate including all the facets of the dimension “personal characteristics”, namely knowledge, skills, and attitudes. As a result, at least one further level of detail could be useful in the existing schemas for describing competences.

b) Measurement scales that represent proficiency levels can be both qualitative and quantitative. Although expecting to use a single, unified measurement scale is not realistic, it would be desirable that at least the values of these scales must be represented in an ordered list, as part of the competence definition schema.

c) The existing approaches to modeling competencies exclude context from their schemas. However, “context” is an important dimension related to competence definition and it should be captured in the competence description (Prins et al, 2008).

At this point, it should be noted that these issues cannot be considered as a list of flaws for HR-XML or IMS RDCEO, since these specifications clearly declare that these areas are outside of their scope.

**Methodology for defining an IEEE LOM Competence-related Application Profile**

As already discussed, educational resources are typically described with metadata using the IEEE Learning Object Metadata (LOM), this does not include metadata elements that are specifically designed for the description of learning resources in terms of their relevance to Competence Development Programmes. In particular, it does not provide metadata elements that would describe the suitability of a given learning resource for supporting learning activities that are designed to meet competence-based educational objectives within competence development programmes.

To this end, we propose to build an IEEE LOM Competence-based Application Profile that can be used for tagging educational resources in a competence-meaningful manner. Hence, based on the key dimensions of competence and the formally defined competence specification models and taking into consideration the CEN/ISSS Learning Technologies Workshop (LTW) guidelines for building application profiles in e-learning (Smith, Van Coillie & Duval, 2006), we can identify possible extensions to the IEEE LOM standard to accommodate competence-related properties.

The process of deriving competence-related metadata consists of the following key steps:

**Step 1: Identify the main dimensions of the concept of competence.** This step aims at identifying and analyzing the various dimensions of the concept competence. The output of this step is detailed in Sampson & Fytros (2008a).
Step 2: Study existing competence description information models. This step aims at analysing the competence description models proposed by international working groups such as IEEE-RCD and HR-XML. The output of this step is the mapping between the elements of the underlying models of these specifications and the identification of aspects that are important for competence modeling taking into consideration the competence dimensions identified in step 1. Details on this matter are given in (Sampson, Karampiperis & Fytros, 2007).

Step 3: Identify the competence relevant characteristics of learning resources. This step is based on the finding of step 2 and it aims to create a schematic representation of competence relevant characteristics for learning resources that will guide our effort to identify possible extensions to the IEEE LOM elements concerning competence related information.

Step 4: Identify suitable IEEE LOM element for accommodating the competence related information. This step examine all the categories and elements of IEEE LOM, in order to find available metadata items that can host the competence relevant characteristics of learning resources.

Step 5: Extending value space or datatype. This step includes the identification of possible extensions required in the value space or datatype of the competence related IEEE LOM elements.

Step 6: Add new sub-elements to the related IEEE LOM element. This step adds new elements and sub-elements to the IEEE LOM existing ones with special attention for avoiding semantic overlaps with other existing elements.

Proposed IEEE LOM Competence-related Application Profile

In this section, we propose an IEEE LOM Competence-related Application Profile that can be used for tagging educational resources in a competence-meaningful manner. To this end, a schematic representation of competence relevant characteristics for learning resources is created to guide our effort for identifying extensions to the IEEE LOM elements concerning competence related information (as described in Step 3). More specifically, the main elements of the schematic representation are as follows:

- **Title**: A short name for the competence that the particular learning object targets at.
- **Description**: A narrative description of the competence that the particular learning object targets at.
- **ProficiencyLevel**: The proficiency level of the competence that the particular learning object targets at. The proficiency level may include a short name and a narrative description. It may also include different types of proficiency level based on the facets of the dimension “personal characteristics” of the term competence, such as “Knowledge”, “Skill” and “Attitude”. Moreover different scales may be used in order to represent proficiency levels. The values of these scales must be represented as an ordered list.
- **Context**: The context of use in which the competence that the learning object targets is referred to.

![Figure 1: Competence characteristics relevant to learning resources](image-url)
The next step (Step 4) is the identification of these elements of IEEE LOM which are considered to be suitable for accommodating competence related information. To this end, we have identified that the IEEE LOM categories which are most related with competence properties of learning resources are: the [5.8 Educational.Difficulty] element and the [9.Classification] Category through the [9.1 Classification.Purpose] element.

The introduced extensions to the IEEE LOM information model are presented in a tabular form for each identified IEEE LOM category. More specifically:

The IEEE LOM [9.Classification] category describes a learning object in relation to a particular classification system (IEEE LOM, 2002). In element [9.1 Classification.Purpose] we can use the “competence” value to state that the purpose of using this particular learning object is to support the attainment of a particular competence. This element already contains a specific vocabulary (for example: prerequisite, accessibility, etc.) that must be updated with the “competence” value for the purpose of (as shown in Table 1).

<table>
<thead>
<tr>
<th>Nr</th>
<th>Name</th>
<th>Explanation</th>
<th>Size</th>
<th>Order</th>
<th>Value Space</th>
<th>Datatype</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Classification</td>
<td>This category describes where this learning object falls within a particular classification system</td>
<td>smallest permitted maximum: 40 items</td>
<td>Unordered</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.1</td>
<td>Purpose</td>
<td>The purpose of classifying this learning object</td>
<td>Unspecified</td>
<td></td>
<td>accessibility Prerequisite … competence Vocabulary (State)</td>
<td>&quot;competence&quot;</td>
<td></td>
</tr>
</tbody>
</table>

The IEEE LOM element [5.8 Educational.Difficulty] describes how difficult it is to work with or through this learning object (IEEE LOM, 2002). In our proposal, this element can be used for the representation of the competence proficiency level that this learning object addresses. However, in accordance to the competence characteristics related to the proficiency level presented in Figure 1, extensions to this element are needed in order to include sub-elements that can describe the values and the scales of a particular competence proficiency level (as shown in Table 2).

<table>
<thead>
<tr>
<th>Nr</th>
<th>Name</th>
<th>Explanation</th>
<th>Size</th>
<th>Order</th>
<th>Value Space</th>
<th>Datatype</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8</td>
<td>Difficulty</td>
<td>The competence proficiency level that this learning object addresses</td>
<td>1</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8.1</td>
<td>Level</td>
<td>There may be multiple instances of this category</td>
<td>3</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8.1.1</td>
<td>Title</td>
<td>Text label of the proficiency level</td>
<td>1</td>
<td>Unspecified</td>
<td>LangString (smallest permitted maximum 1000 char)</td>
<td>Example: &quot;Novice&quot;</td>
<td></td>
</tr>
<tr>
<td>5.8.1.2</td>
<td>Type</td>
<td>The type of</td>
<td>1</td>
<td>Unspecified</td>
<td>Knowledge</td>
<td>Vocabulary</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Element</td>
<td>Description</td>
<td>Skill Attitude</td>
<td>(Enumerated)</td>
<td>&quot;Skill&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------</td>
<td>-------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8.1.3</td>
<td>Description</td>
<td>A human-readable description of the proficiency level</td>
<td>1</td>
<td>Unspecified</td>
<td>&quot;Example: &quot;Adequately performing of the targeted competence&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8.1.4</td>
<td>Value</td>
<td>Rating values for this competence proficiency level</td>
<td>1</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8.1.4.1</td>
<td>NumericValue</td>
<td>NumericValue is the location for quantitative rating scales</td>
<td>1</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 5.8.1.4.1.1 | MinValue | The minimum value of the rating scale | 1 | Unspecified | "Example: "1"
| 5.8.1.4.1.2 | MaxValue | The maximum value of the rating scale | 1 | Unspecified | "Example: "10"
| 5.8.1.4.1.3 | Interval | The increment or step for the relevant scale | 1 | Unspecified | "Example: "1"
| 5.8.1.4.1.4 | Description | A human-readable description of the rating scale | 1 | Unspecified | "Example: "A ten level scale for English language competence" |
| 5.8.1.4.2 | StringValue | StringValue is the location for qualitative rating scales | 1 | Unspecified | |
| 5.8.1.4.2.1 | MinValue | The minimum value of the rating scale | 1 | Unspecified | "Example: "All"
| 5.8.1.4.2.2 | MaxValue | The maximum value of the rating scale | 1 | Unspecified | "Example: "C2"
| 5.8.1.4.2.3 | Description | A human-readable description of the rating scale | 1 | Unspecified | "Example: "A six level scale for Spanish language competence according to Europass" |
| 5.8.1.4.2.4 | Scale | The scale that is used in order to represent | 1 | Unspecified | |
Finally we introduce the addition of a new category with special attention to avoiding semantic overlaps with other existing elements of the information model, namely [10.Competence] category that consists of three main elements namely [10.1 Competence.Title] Element, [10.2 Competence.Description] element and [10.3 Competence.Context] element (as shown in Table 3).

Table 3: Extensions of IEEE LOM new [10. Competence] category

<table>
<thead>
<tr>
<th>Nr</th>
<th>Name</th>
<th>Explanation</th>
<th>Size</th>
<th>Order</th>
<th>Value Space</th>
<th>Datatype</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Competence</td>
<td>This category specifies the competence description</td>
<td>1</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.1</td>
<td>Title</td>
<td>Text label of the competence that the learning object targets</td>
<td>1</td>
<td>Unspecified</td>
<td>LangString</td>
<td>(smallest permitted maximum 1000 char)</td>
<td>&quot;Problem Solving&quot;</td>
</tr>
<tr>
<td>10.2</td>
<td>Description</td>
<td>A human-readable description of the competence that the learning object targets</td>
<td>1</td>
<td>Unspecified</td>
<td>LangString</td>
<td>(smallest permitted maximum 1000 char)</td>
<td>&quot;Identifying problems and implement solutions&quot;</td>
</tr>
<tr>
<td>10.3</td>
<td>Context</td>
<td>The context in which the competence that the learning object addresses is applied</td>
<td>1</td>
<td>Unspecified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.3.1</td>
<td>Title</td>
<td>Text label of the context</td>
<td>1</td>
<td>Unspecified</td>
<td>LangString</td>
<td>(smallest permitted maximum 1000 char)</td>
<td>&quot;Kick-off Meeting&quot;</td>
</tr>
<tr>
<td>10.3.2</td>
<td>Description</td>
<td>A human-readable description of the context</td>
<td>1</td>
<td>Unspecified</td>
<td>LangString</td>
<td>(smallest permitted maximum 1000 char)</td>
<td>&quot;The first meeting with the project team to discuss the role of each team member&quot;</td>
</tr>
</tbody>
</table>

Case Study: Using the proposed IEEE LOM Competence-related Application Profile in a real life project

We now present a full example from a real life case study (namely, the e-Access2Learn Project) to demonstrate and validate our proposal. The e-Access2Learn project (http://www.eaccess2learn.eu) aims to provide tools and services for the development and sharing of accessible eTraining Resources, Activities and Courses that bare the potential to be inter-exchanged between eTraining Platforms and Programmes.
The e-Training Activities and Courses stored to the e-Access2Learn Repository have been tagged with educational metadata following the proposed IEEE LOM Competence-related Application Profile. Therefore, the searching mechanism of the repository, as presented in Figure 2, provide the users the capability to search for e-Training Courses with competence-related searching criteria. More precisely, in Figure 2 we are searching the e-Access2Learn Repository for e-Training Courses, which target the Competence named: “Business Presentation” at the proficiency level of “Novice”. As presented to Figure 3, the repository returned to the user one (1) e-Training Course that matches the searching criteria.

![Figure 2: e-Access2Learn searching mechanism with competence-related searching criteria](image2)

![Figure 3: Search Results of e-Training Courses based on the above searching criteria](image3)
In Figure 4 the full metadata record of the selected e-Training Course is presented. The section “Competence” provides the user detailed description about the Competence that this particular e-Training Course targets at.
Within this framework, the particular instance of the eAccess2Learn IEEE LOM Competence-related Application Profile was validated by a Controlled Method through Replicated Experiment, that is, several instances of the subject (the identified end users of the IEEE LOM AP) performed a number of tasks related with the use of the proposed AP and compared its potential in describing Digital Resources in response to Competence-based Educational Objectives as opposed to the original IEEE LOM Schema. The results of this experiment provided solid indications that the proposed IEEE LOM AP provided the means for tagging educational resource in a competence-meaningful manner.

Conclusions

In lifelong learning, education and training based on attaining competence-based educational objectives appears to become mainstream approach. Within this context, Technology-supported Competence Development Programmes bare the potential of integrating training on-demand in real life work environments and beyond that, they may support the individual lifelong learners in making informed decisions about their continuous personal development during various periods of their lives. The design of Competence Development Programmes assume the clear definition of educational objectives in terms of targeted competences (both type and level), the design of learning activities that are appropriate for the attainment of the defined competences, the efficient selection of suitable educational resources that can implement the learning activities, the accurate definition of the educational setting (tools, delivery means, etc) and the construction of assessment activities that can measure individuals’ current status of the competence level.

Within this framework, the selection of appropriate educational resources based on an accurate and informed estimation of their suitability in accordance to (a) the type of the learning activity to be supported, (b) the educational objectives expressed in terms of competence types and levels, (c) the given educational setting conditions (that is, available tools, delivery means, the presence and roles of tutors and/or other learners), and (d) the individual learner profile (including his/her status of competence). Today, the World Wide Web offers an enormous amount of available resource with potential educational value, both resources that are deliberately designed for support meeting certain educational objective (typically, organised through repositories of learning material), and resources which are produced as the result of real life formal and informal tasks.

Thus, the issue of finding, selecting and assembling digital resources which are suitable for facilitating the attainment of the certain targeted competences is a key issue in technology-supported competence-based lifelong learning. In this paper it was argued that an IEEE LOM Competence Application Profile is needed for tagging educational resources in a competence-meaningful manner for facilitating people and organisations in their search, retrieve, (re)-use and share of appropriate educational resources for Competence Development Programmes. Thus, following the CEN/ISSS LTW guidelines for building application profiles, we presented our proposal for an IEEE LOM Competence Application Profile, based on the competence dimensions identified in our previous work (Sampson & Fytros, 2008a) and the existing competency specification models (IMS RDCEO, HR-XML).

Finally, it would be worthy to mention that the work reported in the paper could be further expoted in facilitating the automatic selection and sequencing of educational resources in a competence-related meaningful manner towards the development of adaptive and personalised learning paths based on competence-related educational objectives. Initial work towards this direction was reported in (Karampiperis & Sampson, 2006) and it could improved by incorporrating competence-related characteristics in the description of the learning resources as proposed in this paper.

Acknowledgements

References


Announcing new **free** resources for online educators!

Participating instructors have experienced an increase in the number of students successfully completing courses, especially learning disabled students with reported student success rates of up to 95%.

www.hancockcollege.edu/improved

- Provides resources for improving distance learning for the learning disabled
- Includes information regarding online learning accommodations for learning disabled students
- Supplies instruction for incorporating Universal Design in Learning to distance learning courses
- Incorporates professional development and training resources for online instructors

This Web site was created in conjunction with the US Department of Education and Allan Hancock College as a demonstration project to ensure students with disabilities receive a high quality education. P333A050026
Outdoor Natural Science Learning with an RFID-Supported Immersive Ubiquitous Learning Environment

Tsung-Yu Liu¹, Tan-Hsu Tan²* and Yu-Ling Chu²

¹Department of Multimedia and Game Science, Lunghua University of Science and Technology, Taoyuan County, Taiwan // Tel: +886-0922488664 // Fax: +886-2-32344001 // joye.liu@msa.hinet.net
²Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan // Tel: +886-2-27712171 // Fax: +886-2-27317187 // thtan@ntut.edu.tw // chu_yuling@tp.edu.tw

*Corresponding author

ABSTRACT
Despite their successful use in many conscientious studies involving outdoor learning applications, mobile learning systems still have certain limitations. For instance, because students cannot obtain real-time, context-aware content in outdoor locations such as historical sites, endangered animal habitats, and geological landscapes, they are unable to search, collect, share, and edit information by using information technology. To address such concerns, this work proposes an environment of ubiquitous learning with educational resources (EULER) based on radio frequency identification (RFID), augmented reality (AR), the Internet, ubiquitous computing, embedded systems, and database technologies. EULER helps teachers deliver lessons on site and cultivate student competency in adopting information technology to improve learning. To evaluate its effectiveness, we used the proposed EULER for natural science learning at the Guandu Nature Park in Taiwan. The participants were elementary school teachers and students. The analytical results revealed that the proposed EULER improves student learning. Moreover, the largely positive feedback from a post-study survey confirms the effectiveness of EULER in supporting outdoor learning and its ability to attract the interest of students.

Keywords
Augmented reality, Collaborative learning, EULER, Outdoor learning, Radio frequency identification (RFID), Ubiquitous learning

Introduction
To enhance the presentation of content in science courses, teachers often use classroom media such as science films or student activities such as experiments. Outdoor teaching is also a recognized method for enhancing learning and is widely used by elementary school teachers in Taiwan. However, the problems encountered in outdoor teaching include a lack of teacher expertise and appropriate teaching schemes. Students often make only quick and casual observations of outdoor conditions and materials, which may not yield sufficient or valuable knowledge. These factors and problems explain the growing interest in information technology applications in this area (Bellotti, Berta, Gloria, & Margarone, 2002; Chen, Kao, & Sheu, 2003; Yatani, Sugimoto, & Kusunoki, 2004).

Digital data and media accessed by mobile electronic devices offer new possibilities in learning. The mobility, flexibility, and instant access provided by modern mobile electronic devices enable students to engage in highly interactive activities at any place or time. Advanced wireless communication technologies enable a new learning model, mobile learning (m-learning). Mobile learning offers a new way to infuse learning into daily life. The advantages of m-learning over e-learning are flexibility, low cost, minimal equipment requirements, and ease of use (Jones, & Jo, 2004). Therefore, many investigations both inside and outside classrooms have assessed the use of mobile devices for enhancing education (Chen, Kao, & Sheu, 2003; Ogata & Yano, 2004; Chen, Kao, & Sheu, 2005).

Moreover, the rapid development of modern computer science and sensor technology has enabled many new ubiquitous computing applications. Ubiquitous computing refers to use of computer systems in everyday environments that enable user interaction at any time (Weiser, 1991; Weiser, 1998; Weiser, Gold, & Brown, 1999). A ubiquitous system provides users with timely information and relevant services by automatically sensing their various context data and by intelligently generating appropriate results (Kwon, Yoo, & Suh, 2005). In ubiquitous learning (u-learning), the ubiquitous computing occurs all around students, regardless of whether they are aware of it or not. Ubiquitous learning provides context-aware information and self-learning opportunities for learners. Thus, it
not only enables students to achieve learning goals anytime and anywhere, it also cultivates their ability to explore new knowledge and solve problems. Ubiquitous learning has the following primary characteristics:

1. Permanency. Learning processes are recorded and stored daily. A learning portfolio not only records the learning processes and achievements, it also records the learning behaviors and environment through a sensor and wireless network.

2. Accessibility. Learners can immediately access information from any location.

3. Immediacy. Learners can immediately access useful information at any time.

4. Interactivity. Learners can interact or communicate with experts, teachers, peers, or the environment. Learning activities can be integrated into the real-world environment and can provide sophisticated, real-time, interactive services.

5. Situation. The learning environment understands the situation of learners by detecting their status via the sensor network.

6. Calmness. The learning devices are quiet, invisible agents that recede into the background of the learning environment.

7. Adaptability. Learners can obtain the right information at the right place in the right way. The system actively adapts the subject content for various devices.

8. Seamlessness. Learners can continue learning processes started elsewhere. The learning environment avoids interruptions in the learning process by adapting ongoing learning to accommodate changes in the learning environment (Chen, Kao, & Sheu, 2003; Ogata, Akamatsu, & Yano, 2005; Li, Zheng, Ogata, & Yano, 2005).

Additionally, we argue that “immersion” is a ninth characteristic of ubiquitous learning. In this context, immersion is defined as the state in which learners experience the same feelings and emotions as in a real world when interacting with virtual objects and environment. Hence, the characteristics of u-learning are permanency, accessibility, immediacy, interactivity, situation, calmness, adaptability, seamlessness, and immersion.

Context-aware systems featuring contextual data retrieval, engaging learning experiences, and improved learning effects have been applied in various learning activities (Cooper, 1993). The term “context-aware” was first defined by Schilit, Adams, and Want (1994), who regarded context in terms of location, identities of nearby people and objects, and the changes to those objects. Dey (2001) defined context as contextual information about an entity, which may be a person, a place, or a physical object. This information is considered relevant to the interaction between a user and an application. Although context-aware ubiquitous learning provides a more flexible, adaptive, and interactive learning experience than mobile learning (Li, Zheng, Ogata, & Yano, 2005), integrating immersion, calmness, seamlessness, and situation into a ubiquitous learning environment is extremely challenging. However, sensors and virtual reality technology are advancing rapidly. Sensor technologies, including IrDA, global positioning system (GPS), Bluetooth, RFID, Zigbee, and wireless local network (WLAN), can provide situated services. However, radio frequency identification (RFID) technology is particularly appropriate because it is inexpensive, highly accurate, very practical, easy to maintain, supports indoor/outdoor learning, and enables situated, calm, and seamless services (Hightower, & Borriello, 2001; Roussos, 2002). An RFID system identifies objects by transmitting internal codes from an RFID chip to a reader. The three main parts of the system are the reader, antenna, and tag. The reader device sends a signal through an antenna to a tag. This signal provides a small amount of power to activate the tag, which then transmits its internal code to the reader (Finkenzeller, 2003).

Conversely, integrating virtual reality (VR) into a ubiquitous learning environment increases immersion and provides a rich learning experience. Virtual reality can be classified as virtual reality, augmented virtuality, augmented reality, and real (Milgram, Takemura, Utsumi, & Kishino, 1994). In virtual reality, the surrounding environments are completely digitalized. In augmented virtuality (AV), real objects are embedded into virtual ones. In augmented reality (AR), digital objects are embedded into the actual environment (Hughes, Stapleton, Hughes, & Smith, 2005). According to Whiteside (2002), an immersive learning environment is effective if it engages the learner holistically — cognitively, emotionally, and even physically — using a combination of design techniques.

Various handheld AR-based learning games have been devised to explore the effectiveness of these technologies for learning. For instance, Wagner and Barakonyi (2003) proposed a handheld AR educational application in which a virtual character teaches users about art history. This study investigates how augmented reality helps create an effective and engaging educational application.
Although the ubiquitous learning systems described above have been successful in outdoor mobile learning, they still have unresolved problems. Specifically, teachers conducting outdoor teaching are unable to present complete information about objects and concepts such as historical relics, habitats of endangered wildlife, and geology landscapes. The following educational issues arise:

- how to achieve museum-like learning experiences in outdoor learning
- how to integrate real learning objects (exhibits) and teaching resources to deliver rich learning experiences to students and thus activate learning motivation and improve learning effects
- how to enable students to search, collect, share, and edit information, so that the goals of independent and collaborative learning can be achieved in outdoor learning
- how to cultivate students’ ability to obtain and share knowledge, solve problems, and creatively collaborate in learning activities?

To address the above issues, this study proposes the environment of ubiquitous learning with educational resources (EULER), which employs RFID, Internet, wireless technology, augmented reality, embedded systems, and database technologies to enhance learning by using mobile learning, immersive learning, and ubiquitous learning with up-to-date technologies. The objective is to provide immersive ubiquitous learning environments for outdoor teaching. Additionally, this study develops interesting learning courses and explores how the proposed EULER affects the learning effectiveness of students. EULER has the following characteristics:

- Accessibility. Learners can use handheld devices to easily access learning resources via the Internet or a wireless local area network (LAN).
- Individuality. Learners can construct knowledge by organizing the learning content and by preparing a learning schedule.
- Immediacy. By using handhelds, learners can immediately access learning resources via the Internet or wireless LAN.
- Interactivity. Learners can interact with peers and instructors and can operate learning objects.
- Social interactivity. Learners can collaboratively complete a common task, discuss a common issue, and share their experiences with each other.
- Connectivity. Learners can exchange information and share files with each other.
- Situation. Learners can gain authentic knowledge of the natural environment.
- Context awareness. Learners can engage in context-aware learning activities in an actual outdoor environment.
- Seamlessness. Ubiquitous computing technology enables a continuous learning process even when the learners change locations.
- Permanency. The learning processes can be recorded in a learning management system and stored permanently.
- Immersion. By interacting with a virtual environment, learners can experience feelings and emotions that are similar to those experienced in the real world.
- Data collection. The system has sufficient multimedia capturing tools to help learners collect images, videos, and audio.

The learning effect of computer-based learning environments continues to be an important research topic. Webster and Hackley (1997) argued that quality, reliability, and ease of use of information technologies influence their effectiveness for learning. Webster and Hackley (1997) concluded that the most effective learning methods should include rich media, student-instructor interaction, and instructors who project positive attitudes. Piccoli, Ahmad, and Ives (2001) found that students with experience using technology and who have positive attitudes about technology tend to perform well in virtual learning environments. Leidner and Jarvenpa (1995) stated that technologies promoting communication and interaction are effective for developing higher-order thinking skills and building conceptual knowledge when following a constructivist or collaborative learning model. Wells (1990) stated that technologies promoting participant communication are best suited for subject matter or course designs that emphasize discussion, brainstorming, problem-solving, collaboration, and reflection. Milheim and Martin (1991) defined three types of control that a learner may have, namely, learning pace control, learning sequence control and learning content control. Learning pace control refers to the speed at which course themes are presented. Learning sequence control refers to the order in which course themes are presented. Learning content control refers to the ability of learners to omit content themes with which they are already familiar. Giving learners some control over course components enhances the effectiveness and efficiency of learning environments according to a literature review by Merrill (1994). Piccoli, Ahmad, and Ives (2001, p. 404) defined interaction as “the degree of contact and
educational exchange among learners and between learners and instructors.” The interactive teaching style is positively related to learning outcomes (Webster & Hackley, 1997), control of the technology and technology attitudes (Leidner & Jarvenpaa, 1995), self-efficacy, and initial motivation. Piccoli, Ahmad, and Ives (2001) stated that students in highly interactive instruction may experience reduced isolation frustration, anxiety, and confusion, and gain positive learning effectiveness. Leidner and Jarvenpaa (1995) reported that the computers can provide learners with exercises similar to those in real-world situations. Given that previous studies asserted that the learning model affects the learning performance of students, this study explores how information technologies, interaction, learner control, and the learning model can improve outdoor natural science learning, based on student feedback by using EULER.

EULER

EULER consists of two subsystems, the mobile interactive learning environment (MOBILE) server for use by teachers and the mobile-tools (m-Tools) for students. Figure 1 shows the structure of EULER. Teachers can give lessons outdoors using a laptop computer with the MOBILE server installed. Teachers can install teaching materials in the mobile content database (MCDB) before class and establish assessments in the mobile assessment database (MADB). The MOBILE server addresses requests from students and responses to students during teaching. The learning activity management (LAM) unit constructs a virtual classroom that supports many learning activities, including bulletin boards, forums, voting, chatting, homework, assessment, and instruction. After completing outdoor teaching, teachers can conduct tests through a mobile assessment management (MAM) unit and analyze student performance. The mobile learning record database (MLRDB) stores student records, including assignment grades, reading times, number of discussions, instances of data collection, and instances of information sharing.

The m-Tools system is installed on a PDA. It provides functions such as m-Loader, m-RFID, m-AR, m-Player, m-Capture, m-Sharer, m-Notes, m-Test, and m-Calendar. The user interface (UI) coordinates each tool and stores the learning records in the mobile database (MDB). For example, when a student using a PDA observes a zoo exhibition with an RFID tag installed, the RFID reader on the PDA identifies the internal code of the tag. The m-RFID then sends this code to the server, which in turn downloads context-aware content to the PDA. The student can then immediately browse the content by using the m-Player. Meanwhile, the m-Player can continuously update the MDB as the learner reads the material. Additionally, the student can use m-Messenger to receive teacher guidance, m-Capture to record videos of animals, and m-Loader to access additional materials from the EULER server. The student can later use m-Notes to compile collected information into articles and then use m-Sharer to send articles to
the team leader via a wireless local area network (WLAN). The team leader can organize these articles, compile them into a team report, and submit the report to the teacher. The collaborative learning activity is then complete. Upon completion, students can employ m-Test to take online tests and evaluate their learning achievement. The m-Calendar can be used to organize personal learning schedules and to remind users about homework deadlines and test events. The m-AR can be used to enable students to see virtual creatures that would rarely appear in the wetland areas during this season.

**Evaluation**

A series of controlled experiments was performed using EULER for outdoor learning by fifth-grade students. After the experiments, a survey was conducted to evaluate EULER in improving student learning motivation and effectiveness.

**Participants**

The participants in this study were students in the Affiliated Experimental Elementary School of Taipei Municipal University of Education, which strongly emphasizes information technology education. Participants included four natural-science teachers and seventy-two students. All participating teachers had at least ten years of experience in computer-assisted instruction. Each student was assigned to either the experimental group or the control group. Each group had six teams of six students.

**Equipment**

The teacher employed a mobile server station, and students used PDAs to perform outdoor learning activities. The mobile server station was comprised of a laptop computer, a wireless access router, a rechargeable battery, and a wooden case (see Fig. 2). The laptop computer was an ASUS M3 with a Windows server 2003, SQL server 2005, and a wireless LAN (IEEE 802.11B/G). The PDA was a wireless-enabled ASUS 716A PDA with Pocket PC 2003, wireless LAN (IEEE 802.11B), Bluetooth, external camera, and external RFID reader. The PDA had a standard memory (SD) slot and a compact flash (CF) slot. For data storage, each PDA had an extra one-gigabyte SD. Additionally, RFID tags were installed in numerous information boards.

![Figure 2. The mobile server station, comprising a laptop, wireless router, and rechargeable battery (right); The learning equipment, including PDA, RFID reader, camera, and RFID tags](image)

**Experimental method**

Table 1 shows the quasi-experimental design of the study adopted for nonequivalent groups. This design required a pre-test and post-test for an experimental group and a control group. The experimental group adopted EULER for learning while the control group adopted traditional methods for learning. All participating teachers had at least ten years of experience in natural science instruction. Tests to evaluate student learning effectiveness were designed by
participant teachers and were updated annually based on instructional requirements. Each test had 20 multiple-choice questions related to the course materials. Under these conditions, the tests had superior validity. Internal consistency of tests and questionnaire in the case study were measured by Cronbach’s alpha coefficient. A Cronbach’s alpha coefficient ranges from 0 to 1, and 0.7 is generally considered the minimum acceptable reliability coefficient (Nunnaly, 1978). Table 2 shows that all Cronbach’s alpha coefficients in this experiment exceeded 0.7, which confirmed the reliability of tests used in the study. Pre-test questions examined prior student knowledge of wetland terminology, geology, and environment as well as related topics.

Post-test questions examined student learning during the course, including the importance of wetlands to human life, and the protection and restoration of wetland ecosystems and wetland creatures. An independent two-sample t-test was performed on the pre-test to determine whether the pre-conditions of two groups significantly differed. In the remaining activities, after the experiments were completed, ANCOVA was adopted to compare group differences in performance on each test with the previous test adopted as the covariance. The statistical data were examined to determine whether the learning performance of two groups significantly differed. Additionally, a statistical analysis was performed to evaluate the performance of the experimental group.

<table>
<thead>
<tr>
<th>Learning phase</th>
<th>Control group</th>
<th>Experimental group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation phase</td>
<td>Introduce the learning topics and goals</td>
<td>Introduce the learning topics and goals</td>
</tr>
<tr>
<td>Treasure-hunt game</td>
<td>Conduct a treasure hunt based on paper materials</td>
<td>Conduct a treasure hunt using EULER</td>
</tr>
<tr>
<td>phase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem-solving</td>
<td>Conduct problem-based learning based on a worksheet</td>
<td>Conduct problem-based learning using the m-Tools</td>
</tr>
<tr>
<td>project phase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immersive learning</td>
<td>Conduct traditional outdoor learning using a text</td>
<td>Conduct immersive learning using the m-AR</td>
</tr>
<tr>
<td>phase</td>
<td>book</td>
<td></td>
</tr>
<tr>
<td>Evaluation phase</td>
<td>Conduct a post-test</td>
<td>Conduct a post-test</td>
</tr>
</tbody>
</table>

Table 2. Internal consistency of the tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>0.84</td>
</tr>
<tr>
<td>Test1</td>
<td>0.76</td>
</tr>
<tr>
<td>Test2</td>
<td>0.80</td>
</tr>
<tr>
<td>Post-test</td>
<td>0.85</td>
</tr>
</tbody>
</table>

To investigate the effect of EULER on student learning, a survey was administered to 36 students after they had completed the experiments. Responses to all questions were on a five-point Likert-scale from 1 for strongly agree to 5 for strongly disagree. Questions were divided into four groups: information technology (Group A), learner control (Group B), interaction (Group C), and learning model (Group D). Table 3 shows the Cronbach’s alpha (α) coefficient of each group. These alpha coefficients exceeded 0.7, which confirmed the internal consistency and reliability of the survey.

<table>
<thead>
<tr>
<th>Group</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. information technology</td>
<td>0.76</td>
</tr>
<tr>
<td>B. learner control</td>
<td>0.79</td>
</tr>
<tr>
<td>C. interaction</td>
<td>0.84</td>
</tr>
<tr>
<td>D. learning model</td>
<td>0.80</td>
</tr>
<tr>
<td>Total</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Course design

Taiwan’s wetlands include many creatures worthy of study. Therefore, a course on Taiwanese ecosystems was chosen for this study. The pedagogic strategy was problem-based learning (PBL) and the designated problem was
“How can the gradually disappearing wetland ecosystems in Taiwan be protected and restored?” The well-documented advantages of PBL over traditional teaching methods include improvements in self-directed learning, self-motivation, problem-solving skills, and knowledge-application skills (Jones, 1996; Stepien & Gallagher, 1993). The learning objectives of this course were as follows: 1) understanding wetland habitats and wildlife; 2) understanding the importance of wetlands; 3) understanding the relationship between wildlife and environment, and 4) understanding the concept and methods of environmental protection for wetland ecological systems. The main learning objectives of this course were to cultivate student problem solving and knowledge construction. The course learning procedures were as follows:

- Phase 1. Preparation phase (week 1)
  Students were divided into control and experimental groups. Teachers administered a pre-test to the two student groups before teaching. The pre-test had 20 multiple-choice questions. Teachers introduced the course topic and learning goals and explained how to use the learning tools.

- Phase 2. Treasure-hunt game phase (week 2)
  The teachers conducted a treasure hunt. The control group students employed a traditional method (using learning sheets) while the experimental group students used PDAs with RFID readers when visiting scenic spots with a treasure map. After answering four questions for each learning unit at the scenic area, the students proceeded to the next scenic area until all five scenic areas had been visited. Students were required to answer 20 multiple-choice questions in this phase for test #1.

- Phase 3. Problem-solving project phase (week 3)
  Students were assigned a problem-solving project based on a defined problem. The students were required to develop hypotheses, find evidence, synthesize knowledge, and explain phenomena. The control group employed the traditional learning method, while the experimental group used EULER to perform activities. Students were required to complete team project reports after school.

- Phase 4. Immersive learning phase (week 4)
  Students participated in an immersive learning activity to learn about wildlife rarely seen in wetlands. The control group students employed a traditional method (reading a book), while the experimental group students adopted the m-AR to see virtual creatures rarely seen in the wetland areas during this season. Teachers gave test #2, consisting of 20 multiple-choice questions, to both groups at the end of this phase.

- Phase 5. Evaluation phase (week 5)
  Teachers graded the project reports and highlighted those that were particularly good. Teachers conducted post-tests and surveys as well as in-depth interviews with teachers and students at the end of this phase.

Learning activities

The learning activities of the control group were based on traditional learning methods. The students who read the textbook and used notebooks to record information were instructed to compile the information on worksheets or in reports. Face-to-face interaction and oral communication between teachers and students were simultaneous. The experimental group, however, performed learning activities by applying EULER. The outdoor teaching was conducted in Guandu Nature Park, a famous wetland in the Taipei area (Wild Bird Society of Taipei, 2007). Student activities included a treasure hunt, a problem-solving project, and an immersive learning activity at Guandu Nature Park in phases 2, 3, and 4, respectively. Teachers used notebooks installed with the MOBILE server to conduct learning activities. Each student performed the learning activities on a mobile learning device with m-Tools installed.

In Phase 1, teachers introduced the course topic and learning goals and explained how to use the learning tools. In phase 2 of the outdoor teaching, the teachers prepared numerous information boards, each with an RFID tag attached. The teachers established the relationship between learning materials and the identification codes of RFID tags then placed the information boards near the corresponding wetland creatures. Each student carried a mobile device equipped with a video camera and an RFID reader when visiting scenic spots using the treasure map. Students approaching a scenic spot could use the learning device to identify the RFID tag attached to the information board. The detected identification code of the RFID tag was then transmitted via WLAN to the MOBILE server used by the teacher. The MOBILE server recognized the location of the students, and then transmitted the context-aware content to their learning devices. After completing a learning unit at one scenic spot, they answered four questions and then continued to the next scenic spot until all scenic spots had been visited. The team that completed the game first was given an award. The students thus accessed context-aware content related to actual wildlife, which enabled context-aware ubiquitous learning. Teachers gave test #1, consisting of 20 multiple-choice questions, to both groups at the
end of this phase. Figure 3 illustrates the scenario at Guandu Nature Park using EULER. Figure 4 depicts examples of learning activities. Figure 5 shows the user interface of the MOBILE server. Figure 6 (a) presents a screenshot of the context-aware content.

In phase 3, students were assigned a team project, “How to restore wetlands.” Students were still unable to propose any feasible method after group discussions and therefore used the m-Messenger to request assistance from the teacher. Under the guidance of the teacher, team members applied m-Capture to capture images or videos of restoration projects currently being implemented in the wetlands and retrieve relevant information from the MOBILE server. This information was then compiled into articles. These articles were then sent to team leaders to create rich team project reports, which were submitted to teachers via WLAN. After all students had completed their outdoor learning activities and submitted their team project reports via m-Share, teachers posted the outstanding reports on the MOBILE server. These outcomes were evaluated by peer groups and teachers.

In phase 4, the immersive learning activity, experimental group students utilized the m-AR to observe virtual creatures that may not normally appear in the wetland areas during this season. For instance, students could see virtual mandarin ducks swimming in a stream. Figure 6 (b) shows a screenshot of virtual creatures. The control group students, however, only saw illustrations of the mandarin ducks in a book. Teachers gave test #2, consisting of 20 multiple-choice questions, to both groups at the end of this phase. In phase 5, the final week, a post-test was given to evaluate the learning performance of all students. A survey was administered to experimental group students after they had completed the course.

Figure 3. Guandu Nature Park scenario employing EULER

Figure 4. Photographs of outdoor teaching activities: (a) students engaging in CSCL; (b) two students engaging in treasure hunt game
Results and discussion

An independent two-sample t-test was performed to analyze the pre-tests of both groups. The difference in average grades between the two groups in the pre-test ($t = -0.21, p > 0.05$) was insignificant, indicating that the prerequisites of the two groups of students were similar. To reduce the effect of student prerequisites on experiments, this study treated the pre-test as a control variable and omitted it. Table 4 shows the results (Sig. > 0.05) of the Levene’s test for equality of variances, which confirmed the homogeneity of variances for the two groups. Additionally, the test results ($p > 0.05$) of between-subject effects, which indicate the assumed homogeneity of regression coefficients for the two groups, were gathered for the remaining tests. Accordingly, ANCOVA analysis was then performed. Table 5 presents the mean grades and standard deviation of evaluations for each test. Table 6 lists the ANCOVA and Cohen $d$ results of the evaluations that were performed for each test. The effect size was used to measure significant differences between the evaluation results for each group. The values 0.2, 0.5, and 0.8 represent small, medium, and large effect sizes, respectively (Thalheimer & Cook, 2002).

In phase 2, the treasure hunt game, the ANCOVA results for test #1 ($F = 18.89, p < 0.05, d = 2.01$) indicate that the average grades of the experimental group exceeded those of the control group by at least 18 points. This statistically significant difference indicated that EULER was effective in improving learning. The interviews suggest that the difference may have been due to the ubiquitous learning activities provided by EULER, which provided a museum-
like experience in outdoor teaching and further increased the student results.

Table 4. Levene’s test for equality of variances and test for equality of regression slopes

<table>
<thead>
<tr>
<th>Test</th>
<th>Test#1</th>
<th>Test#2</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>F</td>
</tr>
<tr>
<td>Levene’s test for equality of variances</td>
<td>2.144</td>
<td>0.148</td>
<td>1.311</td>
</tr>
<tr>
<td>Test for equality of regression slopes</td>
<td>2.984</td>
<td>0.089</td>
<td>0.549</td>
</tr>
</tbody>
</table>

Table 5. Mean grades and SD of evaluations for each test

<table>
<thead>
<tr>
<th>Test</th>
<th>Experimental group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Pre-test</td>
<td>41.4</td>
<td>10.13</td>
</tr>
<tr>
<td>Test 1</td>
<td>80.6</td>
<td>7.10</td>
</tr>
<tr>
<td>Test 2</td>
<td>85.2</td>
<td>9.24</td>
</tr>
<tr>
<td>Post-test</td>
<td>90.1</td>
<td>7.12</td>
</tr>
</tbody>
</table>

In phase 3, the problem-solving project phase, the average grade in the experimental group significantly exceeded (by 21 points) that of the control group ($t = 13.84, p < 0.05$) in the report evaluation, indicating that the proposed EULER and designed learning game improved the collaborative experience and abilities in exploring new knowledge, collecting information, sharing knowledge, constructing knowledge, and working collaboratively. Table 7 shows the independent two-sample $t$-test results for report evaluations.

In phase 4, the immersive learning activity, the ANCOVA results of test #2 ($F = 22.01, p < 0.05, d = 1.26$) indicate that the average grade of the experimental group significantly exceeded (by 11 points) that of the control group. The interviews suggest that this occurred because students could see virtual creatures that may not normally appear in wetland areas. This activity enabled the experimental group students to learn about wetland creatures more effectively than the control group students.

In phase 5, the evaluation phase, the ANCOVA result of post-test ($F = 27.07, p < 0.05, d = 1.80$) revealed that the average grade of the experimental group significantly exceeded (by 13 points) that of the control group in the post-test, showing that the proposed EULER improves learning achievement. The students indicated that EULER provides a ubiquitous learning model that effectively increased their interest and learning motivation.

Table 6. ANCOVA and effect size ($d$) results for two group tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Variance source</th>
<th>SS</th>
<th>MS</th>
<th>$F$</th>
<th>Effect size ($d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>Pre-test error</td>
<td>921.88</td>
<td>3367.67</td>
<td>18.89*</td>
<td>2.01</td>
</tr>
<tr>
<td></td>
<td>Test#1 error</td>
<td>1911.76</td>
<td>5992.13</td>
<td>22.01*</td>
<td>1.26</td>
</tr>
<tr>
<td>Test 2</td>
<td>Test#2 error</td>
<td>1399.97</td>
<td>3568.47</td>
<td>27.07*</td>
<td>1.80</td>
</tr>
</tbody>
</table>

* $p < 0.05$

Table 7. Independent two-sample $t$-test results for a project report

<table>
<thead>
<tr>
<th></th>
<th>Experimental group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t$</td>
<td>Sig.</td>
</tr>
<tr>
<td>Project report</td>
<td>13.84</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

* $p < 0.05$

Additionally, according to Table 8, the average post-test grade in the experimental group significantly exceeded that in the pre-test ($t = 40.42, p < 0.01$). This finding also demonstrates the effectiveness of EULER. Figure 7 plots the bar charts of the grades of the two groups. The average grade of the experimental group gradually increased and was
higher than that of the control group, further indicating that the experimental group learned more effectively than the control group did. This finding also indicates the effectiveness of EULER.

| Table 8. Dependent $t$-test results of pre-test and post-test for experimental group |
|---------------------------------|-------|---|--------|
| Pre-test and post-test pair      | $t$   | Sig. | d.f.   | M.D.  |
|                                 | 40.42 | 0.000* | 35     | 47.98 |

Test value = $41.40$; *$p < 0.01$

Figure 7. Bar chart of grades of two student groups, indicating that the experimental group learned more effectively than the control group did

A questionnaire was administered to the experimental group at the end of the learning activity. All 36 of the questionnaires distributed were valid, which was a 100 percent response rate. Table 9 lists the survey results, based on a 5-point Likert scale. The responses to item A1 indicated that most students thought that the m-Tools were easy to use ($m = 3.83$). One student stated, “I wanted to know whether crabs all move sideways. I found that one of the species called the soldier crab (*Mictyris brevidactylus*) is an exception. Using these ubiquitous learning tools, I could acquire new knowledge easily.” The responses to item A2 indicated that the system functions were convenient and sufficient for learning ($m = 3.86$). Most students considered PDAs to be novel devices, and opined that such learning experiences would not be possible without the handheld devices. Students also commented that the m-Tools had many functions with user-friendly interfaces that could assist them in completing the desired learning activities. One student stated, “Whenever I observed anything that I found interesting, I would use the video-capturing tool (m-Capture) to record the scenes and save them in the PDA memory card.” The responses to item A3 indicated that most students effectively used the m-Tools to collect, share, and analyze data, and further complete reports ($m = 4.02$). Many students indicated that using modern devices when learning outdoors is a diverting experience. They reasoned that m-Tools develop their abilities to use information technology to a degree not possible with traditional methods. One student stated, “I grew up in a big city. Except for occasionally visiting the zoo, I seldom had the chance to be exposed to wildlife. Most of my knowledge was acquired from textbooks. This outdoor learning activity allowed me to watch wild creatures but also deepened my understanding about their living environment and activities.” Result A3 also replied to the issue I4, reflecting the ability of EULER to cultivate student competence in acquiring and sharing knowledge, solving problems, and becoming creative in learning. The responses to item A4 indicated that most students thought that the screen size of the PDA was too small ($m = 3.02$). A larger screen would enable them to read the content easily. The responses to item A5 indicated that most students considered the audio quality of PDA inadequate ($m = 3.28$), such that they could not listen to materials clearly.

Item B1 indicates that students gained much more knowledge by using m-RFID than by using worksheets ($m = 4.32$). One student stated, “I could not only read webpage-based materials at any time, but could also watch biological creatures closely. I could verify what I had learned from textbooks with the knowledge I had gained in real life.” Item B2 indicates that students could collect more information by using m-Loader and m-Capture ($m = 4.12$).
One student stated, “I used m-Capture to capture some valuable photos and films. Many of my classmates were very jealous of me after they saw the photos and films.” Item B3 indicate that most students thought that the m-Sharer could help them share information with classmates \((m = 4.02)\). Item B4 indicates that most students thought that m-Notes could help them complete common projects collaboratively \((m = 4.06)\). One student stated, “Our group used m-Share to transfer the animal data that we gathered to the leader. After the leader prepared a report, he shared it with all the group members. Later, using the discussion tool, we discussed how to improve the report, where to add more text, and where to add new pictures. Through many discussions and revisions, we jointly completed the final report.” Results B1, B2, B3, and B4 also replied to the issue I3, reflecting the ability of EULER to enable students to search, collect, share, and edit information further to improve learning.

The responses to item C1 indicated that most students thought that using the m-Capture could increase interaction with natural environment \((m = 4.22)\). One student stated, “I used the image-capturing tool (m-Capture) to take snapshots of animals that I observed. These picturesque records were authentic knowledge that I could not possibly learn from textbooks.” Responses to item C2 indicated that most students thought that using the m-Messager increased interaction with teachers \((m = 4.13)\). Responses to item C3 indicated that most students thought that employing the m-Loader and m-Notes increased interaction with computers \((m = 4.22)\). Many students commented that they were highly motivated to use modern devices such as PDAs. Responses to item C4 indicated that most students thought using the m-Sharer could raise interaction between classmates \((m = 4.17)\). Results C1, C2, C3, and C4 also replied to the issue I2, illustrating the ability of EULER to motivate students in learning and improve learning.

### Table 9. Summary of survey results from 36 students (5-point Likert scale)

<table>
<thead>
<tr>
<th>Group</th>
<th>Item</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Information technology</td>
<td>A1. The user interface of the m-Tools is friendly.</td>
<td>3.83</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>A2. The system functions of the m-Tools are convenient.</td>
<td>3.86</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>A3. I am familiar with using m-Tools to collect, share, and analyze data, and to further compile them to a report.</td>
<td>4.02</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>A4. The PDA display is sufficiently clear for reading learning content.</td>
<td>3.02</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>A5. The PDA audio is sufficiently clear for listening to learning content.</td>
<td>3.28</td>
<td>0.98</td>
</tr>
<tr>
<td>B. Learner control</td>
<td>B1. Applying the m-RFID to outdoor learning helps me gain real knowledge.</td>
<td>4.32</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>B2. Applying the m-Loader and m-Capture to outdoor learning helps our team in collect information.</td>
<td>4.12</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>B3. Applying the m-Sharer to outdoor learning helps our team to share information.</td>
<td>4.02</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>B4. Using m-Notes for outdoor learning helps our team complete assigned project collaboratively.</td>
<td>4.06</td>
<td>0.84</td>
</tr>
<tr>
<td>C. Interaction</td>
<td>C1. Using the m-Capture increases interaction with the natural environment.</td>
<td>4.22</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>C2. Using the m-Messager increases interaction with teachers.</td>
<td>4.13</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>C3. Using the m-Loader and m-Notes increases interaction with computers.</td>
<td>4.22</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>C4. Using the m-Sharer increases interaction with classmates.</td>
<td>4.17</td>
<td>0.80</td>
</tr>
<tr>
<td>D. Learning model</td>
<td>D1. I enjoy the problem-solving learning activity because it enables me to understand the importance of wetland ecology protection.</td>
<td>4.12</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>D2. I enjoy the treasure-hunt game because it makes outdoor learning interesting.</td>
<td>4.25</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>D3. I enjoy the immersive learning activity because it enables me to watch virtual creatures in the actual natural environment and to learn about wetland wildlife.</td>
<td>4.33</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>D4. I enjoy the RFID-based ubiquitous learning environment, which is similar to a natural museum.</td>
<td>4.22</td>
<td>0.74</td>
</tr>
</tbody>
</table>

\(d.f. = 35\)

Responses to item D1 indicated that most students enjoyed the problem-solving activities and understood the importance of wetland ecology protection \((m = 4.12)\). One student replied that household wastewater, industrial wastewater, and agricultural chemicals all eventually flow into rivers. He thought wetlands became damaged because river water passes through them before running into the ocean. Responses to item D2 indicated that most students
found that the multimedia learning materials and the on-site treasure hunt game added interest to the learning process \( (m = 4.25) \). Another student viewed the treasure-hunt game as new and interesting. She said, “I saw fiddler crabs waving fiddle claws and waterbirds looking for foods beside the pond. I could also reach out my hands to touch the Kandelia candel. In this activity, I could not only read webpage-based materials at any time, but also have a close watch on biological creatures. I could verify what I had learned from textbooks with the knowledge I gain in real life.” Responses to item D3 indicate that most students observed many creatures that they had never seen before \( (m = 4.33) \). Many commented that they enjoyed watching virtual creatures on PDAs. One of them stated, “Through the m-AR function of the PDA, I could see a virtual Uca lacteal waving its fiddle claws on the swamp.” Another student stated, “Through the m-AR, I could see virtual creatures that I did not see in the wetland.” Responses to item D4 show that most students regarded EULER as a natural science museum \( (m = 4.22) \). One student stated, “When I walked into a learning area, the RFID sensor device on my PDA could detect my location and tell the PDA to download relevant materials for me. I felt like I was visiting a museum.” Result D4 also replied to the issue I1, illustrating the ability of students to attain museum-like learning experiences in this outdoor learning environment. Knowledge impressed students most favorably through actual, repetitive, and persistent operation of mobile devices. Therefore, student learning effect can be improved.

The interviews indicated the following reasons for the improvement:

1) EULER supports convenient and flexible functions and provides diverse educational applications for teachers.
2) EULER provides plentiful learning resources; students can learn flexibly and independently at any time and place.
3) EULER affords amusing learning activities, which increase learning motivation.
4) The m-tools help students to solve problems and stimulate their creativity through computer-supported collaborative learning (CSCL) activities.

Conclusions

This work presents an environment of ubiquitous learning with educational resources (EULER) scheme based on RFID sensor technology, the Internet, ubiquitous computing, embedded systems, and database technologies to help teachers develop diverse learning applications anytime and anywhere — even in locations without information presentation capacity. The proposed EULER consists of two subsystems: the EULER server installed on a notebook computer for use by teacher, and the mobile learning tools installed on a mobile device for application by a student. The EULER server helps teachers conduct outdoor learning activities involving various educational strategies, including collaborative learning, problem-based learning, and context-aware learning. Additionally, the mobile learning tools support student learning anytime and anywhere. A case study of natural science learning applying RFID technology was performed in classrooms and at the Guandu Nature Park with a learning topic entitled “Taiwan wetland ecosystems.” The participants were four elementary school teachers and 72 fifth-grade students. Problem-based learning, computer-supported collaborative learning, and ubiquitous learning pedagogic strategies were adopted, and a series of formative learning activities were employed in five-week courses.

Test results obtained from learning activities revealed that the experimental group achieved significantly more learning improvement than the control group did. A survey was administered following all learning activities. The results indicated that most students found EULER easy to use, and most believed it could assist their learning. Further, the interview results reveal that, compared to traditional learning methods, the proposed EULER not only increases the motivation of students to learn and improves the effectiveness of learning, but it also improves student creativity and the ability to explore and absorb new knowledge and solve problems. The analysis of students’ reports and interviews concludes that the destruction of wetlands caused by industrial pollution, solid waste, and engineering construction. We recommend that the government promote the concept of environmental protection and administer more artificial restoration projects.

This work demonstrates that RFID and AR technologies provide museum-like learning experiences in context-aware, immersive, and ubiquitous learning activities. Moreover, educators can benefit from the inclusion of educational objectives of acquisition of authentic knowledge, development of the sense of cooperation, training of problem-solving ability, improvement of learning motivation, stimulation of imagination, and creativity in future learning activities. Through these new learning activities, students’ experiences in using technology to facilitate their learning can definitely affect their future learning. Hopefully, using EULER to present well-designed immersive ubiquitous
learning can provide enjoyable and effective outdoor learning experiences.

In future work, efforts are underway to describe not only the educational affordances of EULER towards outdoor, natural science learning, but also the detailed process of interaction, collaboration, and learning by using EULER. Additionally, one new project is underway to employ new sensor networks, physical interaction, artificial intelligence, and ubiquitous AR technologies in the learning environment. Artificial intelligence and AR technologies enable students to virtually observe the behavior of creatures that would rarely appear in the wetlands and even creatures that no longer inhabit the earth. Students can interact with virtual creatures by using sensor networking and physical interaction technologies. This feature allows students to envisage the impact of human behavior on the environment and on other creatures and helps them to appreciate the need for environmental protection and understand how to live with nature.

Acknowledgments

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC94-2520-S-027-002, NSC95-2520-S-027-002, and NSC97-2511-S-262-002.

References


An Investigation into E-Tool Use for Formative Assignment Assessment – Status and Recommendations

Eva Heinrich¹, John Milne² and Maurice Moore³
¹School of Engineering and Advanced Technology // ²Centre for Academic Development and e-Learning // Massey University, Palmerston North, New Zealand // E.Heinrich@massey.ac.nz // J.D.Milne@massey.ac.nz
³Appreciative Developments, Wellington, New Zealand // maurice@mjm.co.nz

ABSTRACT
This article reports on a comprehensive study, investigating the use of e-tools for formative assignment assessment. The study conducted a large-scale literature review and interviews with 90 academics at five New Zealand tertiary institutions. The focus of the study was on formative assessment provided in assignments, an area in which educational theories on importance for learning and on best practice are well established. Yet, research shows that these theories have failed to influence tertiary practice sufficiently to ensure the best possible outcomes for student learning. The interview data collected were analysed according to motivation, needs and benefits as perceived by lecturers and educational principles as derived from the literature. Both the literature and the interview analysis showed a strong positive impact from using e-tools for the management and marking of assignments. Widely available e-learning and general purpose computer tools can be used effectively to support lecturers in dealing with assignments. Specialised assignment tools provide opportunities for even greater gains. The use of e-tools means time savings for the lecturers and enables quality gains in the assignment process. E-tools have the potential to assist with uptake of the well-established educational theories. This article summarises important aspects of the literature review and the interview analysis conducted as part of the study. It provides recommendations in the form of suggestions on which tools are suitable for specific steps in the assignment process and alerts the reader to the website where there are user profiles supported by case studies and an action plan. All these recommendations are targeted at encouraging and enabling lecturers to employ e-tools for the management and marking of assignments.

Keywords
Assignments, Assessment, E-learning tools

Introduction
Formative assessment provided in assignments is of highest value in student learning. Assignments require students to construct and formulate their own thoughts and, depending on discipline, present these thoughts in the form of essays, models, calculations, or reports. Educational theories around the importance of assignments and best practice are well established. The questions that led to the research project reported here were how e-tools can help lecturers use good practice in formative assessment with assignments, to which degree the use of such tools is currently happening, and how further uptake can be encouraged.

To answer these questions, a comprehensive literature review and interviews with 90 lecturers at five New Zealand tertiary institutions were undertaken. General and specialised e-learning tools suitable for supporting aspects around assignment processes were investigated. This article summarises important aspects of the literature review and the interviews. It focuses on recommendations directed towards individual lecturers who are striving for gains in efficiency and quality by suggesting ways of integrating e-tools into their assignment practice. The full findings and details are available at the project website, http://etools.massey.ac.nz.

Summary of the literature review
The literature review for the project had two main aims. The first was to establish the characteristics and importance of formative assessment with assignments. The second aim was to identify electronic techniques, tools, and approaches for this type of assessment, to discuss what was known about their effectiveness, and to uncover factors influencing uptake.

For the first aim, key articles and books that have shaped current theory on formative assessment of student work have been identified and analysed. For the second aim, a comprehensive search of e-learning literature from 2001 to
2006 was conducted. A list of about 100 journals and conferences relevant to assessment and e-learning were identified. From these a subset of close to 40 sources was selected, based on journal impact factor, reputation, and relevance to the study. Within these sources over 580 journal editions and conference proceedings were searched.

Assessment with essay-type assignments

Formative assessment with assignments is part of coursework. Coursework implies both a formative component, feedback to facilitate learning, and a summative component, the marks the students receive for their efforts. The importance of this type of assessment lies in the challenges it poses to students to formulate their own thoughts and construct their own answers, linking to higher-level learning outcomes. The authors of educational literature typically use the term “essays” to refer to such coursework. While many disciplines do not use the term “essay,” the challenges and benefits noted apply to a wider body of student work. For example, in business studies one might talk about a report; in the sciences students might describe the planning and analysis of laboratory work; in a software development context students might be required to elicit requirements and formulate models. All this type of work poses the challenges of formulating own thoughts and constructing answers. The next two paragraphs, reporting directly from the literature, use the term “essays.” The later sections of this article refer to “essay-type” assessment to alert to the wider context.

Essay assignment questions provide students with freedom of response and require students to apply their knowledge and skills in organizing, integrating, and evaluating information (Linn & Miller, 2005; Nitko, 2004). Essays target knowledge, understanding, and proficiency in reasoning (Stiggins, 2005). Essays aim at the highest level of learning outcomes by emphasising the integration and application of ideas (Gronlund, 2006). In contrast to multiple-choice examinations, the answering of essay questions prompts students to adopt deep learning approaches (Struyven, Dochy, & Janssens, 2005).

Essay assignments require the students to formulate their own responses. Writing requires thinking and reflection to be able to communicate knowledge in clear, plausible, and effective ways (Tynjala, Mason, & Lonka, 2001). Data and ideas need to be organized and integrated (Linn & Miller, 2005). Writing is integral to the process of learning (Goodfellow & Lea, 2005). It plays an important part in language development and knowledge construction (Lindblom-Ylanne & Pihlajamaki, 2003). With these characteristics, writing plays an important part in achieving higher-level learning outcomes.

E-learning tools

Areas in which e-learning tools and approaches can support essay-type assessment are the management of assignment submission, storage and return, the assistance in providing individualised feedback to students, the coordination of marking teams, and the communication with students (Stephens, Sargent, & Brew, 2001).

Various types of software can be used to support essay-type assessment. Learning management systems provide base-level support, largely for the management of assignments. General purpose tools such as word-processing systems can be adapted for tasks like providing feedback. Specialised assessment software is available. It falls into four major groupings of which only one is relevant to the formative assignment assessment considered here. The first set is specialist software for the specification and automated marking of restricted-response item assessment like multiple-choice tests. These tools are not relevant in the context of this research as they target a different assessment paradigm in which students select from pre-defined answers and the correctness of an answer is pre-determined. While valuable in their own right, these tools are not included as they do not match the “essay-type” nature of assessment considered here.

The second set is specialist software that attempts the automated assessment of free-form student writing. Several articles in the literature report a relatively good reliability of these systems. The main reasons why these systems are not considered further in this research are as follows: the need for very large essay sample sizes to calibrate the automated marking; restrictions in providing meaningful individualised feedback to students; and the limitations of only being able to analyse text-based student work.
The third set of specialist software comprises tools that have been developed for automated feedback in specific domains. For example, certain aspects of computer programs, such as correctness, style, documentation, or efficiency, can be analysed automatically and formative feedback can be generated. This set of tools is not included in this work for two reasons. First, computer programs use strict syntax rules and do not fit the idea of “free-form” as followed in this article. Second, computer programming is a highly specialised area, whereas the work in this research aims at identifying e-learning solutions with wide applicability.

The fourth relevant set of specialist software for assessment is comprised of tools for marking assignments. These tools build on educational assessment theories. They offer support for the management of assignments and, most importantly, provide features conducive to high-quality formative marking (Stevens & Jamieson, 2002; Blayney & Freeman, 2004). The tools integrate scoring rubrics that allow markers to provide feedback at a detailed level and in summary. The general philosophy of these tools is to liberate the marker from mundane but necessary administrative tasks and to create an environment that facilitates quality feedback, enhancing student learning. Examples of these tools are Turnitin and Grademark (http://www.turnitin.com), Electronic Feedback Software (http://www.ljmu.ac.uk/cis/software/feedback.asp), Questionmark Perception (http://www.questionmark.com), WebCTConnect and MarkTool (http://www-ist.massey.ac.nz/marktool/), Markin (http://www.cict.co.uk/software/markin/index.htm), and Markers Assistant (http://www.surreal.com.au).

While the software in this last set shows a lot of promise for essay-type assessment, it is not widely used. Uptake is localised to institutions where researchers have developed these tools and even at these institutions only early adopters are involved (Sim, Holifield, & Brown, 2004). No reliable studies are available that would show numbers of users across the tertiary sector. The developers of the tools report on the use of their tools (Baillie-de Byl, 2004; Denton, 2003; Wells, 2006), but this provides only examples of localised use. The authors of this article have done an informal evaluation of a specialist tool available at their institution, which has indicated a low uptake by only a few percent of instructors. This needs to be seen in light of the low uptake of the assignment tool that is part of the learning management system at this institution, which is the required first step for using the specialist tool (http://www-ist.massey.ac.nz/marktool/, see Survey Results).

Stakeholder groups

The main groups of stakeholders concerned with assessment are students, lecturers and institutional support, and management staff. Students in general are concerned with the fairness of assessment and the level of feedback they receive. Students are open to new forms of assessment and to the use of e-learning tools in general. Based on the high stakes involved in assessment, students are not prepared to tolerate any form of malfunction of an assessment system (Campton & Young, 2005; Nesbit & Burton, 2006).

Lecturers live in a complex network of their own, student, and institutional needs. Their considerations regarding uptake of e-tools for assessment are influenced by opinions and experiences of colleagues, familiarity with technology, opportunities for efficiency gains, preparedness in modifying their approaches to teaching and learning, and institutional support. The use of e-tools for assessment can increase the awareness of staff regarding assessment approaches (Aller et al., 2005) and can positively influence the link between assessment and teaching and learning (Buzzetto-More & Alade, 2006). A crucial factor in the evaluation of the effectiveness of an e-tool lies in the flexibility of staff. Staff who are willing to adjust their assessment practice can benefit greatly from such an electronic system (Jones, Cranston, Behrens, & Jamieson, 2005). The selection of e-tools for assessment must be guided by the pedagogical design of the assessment. It is up to staff to do the thinking, take control, and match tool use to their assessment approach and learning goals (Freeman & McKenzie, 2002). The primary focus needs to be on pedagogical improvements while productivity gains can only be hoped for in the medium to long term (Warburton, 2006).

Institutional support and management is of critical importance. Institutions must be instrumental in making the appropriate e-tools available, guiding and protecting staff with suitable policies, providing comprehensive training both on instructional and technical levels, and ensuring ongoing support (Buzzetto-More & Alade, 2006; Freeman & McKenzie, 2002; Sim, Holifield, & Brown, 2004; Warburton, 2006).
Importance of e-learning technology for assessment

There are a variety of ways in which e-learning can facilitate assessment in general and essay-type assessment in particular. Freeman and McKenzie (2002) advocate the use of web-based systems to improve students’ learning of teamwork skills. They state that such systems reduce problems with teamwork, especially in the context of large classes. Stevens and Jamieson (2002) talk about the use of e-tools to improve marking quality and feedback. Blayney and Freeman (2004) list a number of ways in which computing support can help with providing feedback. They talk about supporting human markers in commenting on supply items, about polling student opinions in the classroom, and about getting insight into student understanding through quizzes and tests. Plimmer and Mason (2006) as well as Edwards and colleagues (2002) emphasise the advantages of electronic submission and handling of assignments. They list issues like the ease in collection of student work, the removal of geographic limitations, the reduced risk of lost work, the time and resource savings if printing is not required, the long-term availability based on the ease of storage of electronic artefacts, and the efficient return of marked student work. A number of authors point out the opportunities for anonymous participation and marking that arise from e-assessment approaches (Davies, 2002; Downton, 2006; Edwards, Fernandez, Milionis, & Williamson, 2002). Student identities can be stored in a database and the artefact that the marker sees can be kept anonymous, which is one of the recommendations for addressing the challenges in the reliability of essay marking (Gronlund, 2006; Lambert & Lines, 2000). Peer environments can be designed in which students can participate anonymously, with only the assessing teacher knowing the students’ identity. New assessment designs are possible in which students learn from repositories of marked example assignments (Heinrich, 2004; Heinrich & Lu, 2005). These designs include self- and peer-assessment and scaffolding of student learning by only partially releasing feedback data.

It is essential to use technology for the right pedagogical reasons. The use of technology for its own sake will not improve educational assessment (Committee on the Foundations of Assessment, 2001). The use of technology must be well thought-out, and assessment must be integrated into learning (Blayney & Freeman, 2004). Students should be informed of the reasons for using e-learning systems and the relation of these systems to learning design (Freeman & McKenzie, 2002). New approaches to assessment in higher education can be informed by e-learning technology. These approaches should emphasise the role of technology in supporting the human assessor, whose role is vital, especially for formative assessment (McGuire, 2005). For any form of assessment, validity and reliability are important measures. In an e-learning context new forms of assessment can and should be explored to achieve these measures, and this process for improvement should be driven by all stakeholders involved in education (Wijekumar, Ferguson, & Wagoner, 2006).

In summary it can be said that there is a lot of opportunity for using e-learning technologies and approaches to improve assessment. There is also consensus in the literature that these opportunities are not being taking advantage of and that not enough effort is put into exploring the use of e-learning for assessment processes and tasks (Baillie-de Byl, 2004; Blayney & Freeman, 2004; O’Reilly, 2005; Wells, 2006).

Summary of interviews

The survey comprised interviews with staff who use e-tools to help receive, mark, or return assignments. The sampling strategy was purposive. The selection deliberately looked for staff using e-tools for the marking and management of assignments. The aim was to examine the practice of the leaders in the field so that the value of e-tool use could be assessed and recommendations be formulated on good practice. This study did not attempt to identify how many lecturers use e-tools. The selection process took care to include lecturers who use a diversity of assignments and educational contexts covering a range of class sizes, subjects, and educational levels. Key contacts at each institution selected staff they judged to best fit the sample criteria.

The interviews were semi-structured and guided by the questions about the context of the assessment, the use of e-tools, views on advantages and disadvantages of tool use, and factors determining tool use. Information was collected on demographics such as years of teaching, class size, level of degree, subject (based on New Zealand standard classification of education [Ministry of Education, 2007]), course type (campus or distance), and institution. The participants were predominately experienced lecturers. About half of the sample had taught in tertiary education for more than ten years, 40 percent had taught between three and ten years, while fewer than 10 percent had taught
less than three years. The courses the participants discussed were an even mix of campus-based and distance courses. Forty percent of courses were campus based, 40 percent were distance, and the remaining 20 percent were either both campus-based and distance or the participant discussed both modes using different course examples.

Interview analysis directions

The following main analysis themes and sub-themes were selected:

- Benefits. Administration, reducing geographic limitations, time and resource savings, group facilitation, archives, plagiarism, and other issues
- Motivation of lecturer. Institutional criteria, benefits for students, pedagogy, productivity gains, preventing plagiarism, tools being a necessity or reasons for using a specific tool
- Needs of lecturer. General needs of lecturer, specific needs of the lecturer, technology requirements, and student needs
- Educational aspects of feedback. Clarify goals and criteria, self-assessment, feedback, dialogue, motivation, closing the gap, shaping teaching, validity.

The analysis on educational aspects of feedback was guided by Nicol and Macfarlane-Dick’s work on the seven principles of good feedback (2006) and the need for validity of assessment. To guide the analysis of the interview transcripts, the principles were complemented by thoughts on practice, expressing how the more theoretical concepts could manifest themselves in teaching.

The theme-based coding was applied fully to 30 interview transcripts. The majority of the 30 interview transcripts were chosen based on the following four criteria:

- Well-defined need matched by the use of technology. The needs of the students and the course were met by the technological innovation employed. An important point was that there was good use of e-learning tools.
- Good feedback and assessment practice. Evidence that the lecturer was adhering to pedagogical principles or the seven principles of good feedback practice outlined by Nicol and Macfarlane-Dick (2006).
- Evidence of innovation in the area of using e-tools for assessment. For example, did the lecturer create his or her own system to meet the needs of the students?
- Clear interview that added something new. A comparison of interviews on the range of tools and lecturers from different institutions helped the selection process.

In addition, all interviews were analysed based on a list of issues, focusing on the key principles of good assessment, the steps in the assignment process, and the extent of tool use. The issue analysis is not reported here but can be found in the reports on the project website (http://etools.massey.ac.nz/).

Selected analysis results

From the rich set of analysis results only a few can be reported here. This section presents some of the results from the analysis on benefits, motivation of lecturer, needs of lecturer, and educational aspects of feedback.

Benefits

In the benefit analysis, the benefit that was referred to most often was e-tools to assist with administration issues. Following this were the major benefits of time and resource saving, facilitating group activity, the removal of geographic limitations, improving the quality of marking and feedback, and having an archive of student work. A practical benefit of using e-learning tools mentioned most often by the lecturers was reducing effort and time they spent on administration.

R46: One of the reasons I really started to do this was because with distance papers I felt a little bit like the Harry Potter movie when the letters [arrived] and [were] flying everywhere and I felt that I had essays coming under the door, I had essays in my mailbox, essays by post, essays on the computer, essays on digital, and it was everywhere and I just thought right submit them digitally and now it’s just via the digital...
drop box.

**Motivation**

One of the strong motivational factors for using e-learning tools for assessment was based on good institutional support. A common theme that emerged from the interviews was that lecturers were motivated to use the learning management system the university supported. A number of lecturers suggested they were “forced” or had little choice but to use the system supported by the university, and there were no options for using alternative systems. Another group described this as a form of protection for lecturers as the responsibility did not rest with them. If something went wrong with the system, assistance was readily available.

**Needs**

A variety of different needs of lecturers arose when discussing feedback and the use of e-tools. These were grouped into five main themes. First, needs that relate to course context, such as class size and availability of tutors, were identified. Second, a need for institutional support emerged. Third, a tool’s ease of use was recognised. The fourth need involved specific technological needs or improvements related to individual tools. Finally, student needs were also identified by the lecturers. The following quote expresses the frustration of a lecturer who is required to spend too much time on administrative tasks and who recognises this as a detriment in assisting student learning.

R36: I think the frustration for me is I would like to be able to give much fuller feedback. I mean one of the big deals for me is that I teach really big courses, so I actually don’t really do a lot of my own marking, and sometimes I feel like I would walk over broken glass to be able to teach two courses with 20 students in them so I could do all my own marking and take my time over it. So I spend a lot of time checking marking and arranging marking rather than marking, and that is so frustrating.

The need for institutional support is clearly recognised in the literature. Essential to the successful adoption of e-tools for assessment is a strong support network and support from the institution (Freeman & McKenzie, 2002). It has been reported that staff often experience a lack of support and disinterest from colleagues when using new e-tools (O’Reilly, 2005). Not only do staff need technical support on how to use a tool and how to mark using the tool, but also the environment around the staff member needs to support use of the tool. A number of reports suggest that the introduction of e-tools for assessment requires a holistic institutional approach (Downton, 2006; Kennedy, Webster, Benson, James, & Bailey, 2002; Kenny, 2002). In agreement with this, a number of lecturers mentioned the need for more support from within an institution about how to use e-tools and for help with problems, as illustrated by the extract below:

R38: I think that they should do some research, provide some guidelines, and provide some suggestions, what packages are out there. When we were looking into Turnitin I went, this should be something that’s funded, there should be a central investigation as to what a good tool is or what some good tools are and then it should either be provided centrally... rather than every single lecturer who’s interested and are struggling along and doing it on their own. Why can’t they put together ways of marking electronically?

Looking at needs related to technology, a number of devices have been identified as having possible uses in the provision of feedback to students. Audio comments and visual feedback were needs identified by a number of lecturers from the interviews as a whole. It was thought that being able to give audio or visual feedback to students could save time in giving feedback and make feedback clearer to students. One lecturer’s views about media-rich feedback are below:

R5: There’s clear opportunity for tools to help in the automation of this, and my feeling is that a more media-interactive or more media-rich feedback with a combination of textual, audio, and some video conferencing or something where I can actually point at things would be really, really nice. If that was so easy that it would be no more different than me spending ten minutes pointing and talking to them, then that would be fabulous because it’s taking me that long to type things in.
Other needs that were mentioned depended on the particular tool that the lecturer used. Some lecturers were interested in tools that could help in the provision of more structured feedback, such as a database of standard comments that could be chosen from while marking. Some highlighted limitations of the tools that were used. For example, some lecturers complained that learning management systems were too clunky to use, or that they were not integrated with the centralised university system, and they wished for improvements in this area.

Addressing the needs of students, one lecturer noted that once a course is finished students should be able to access the archive of their work and feedback that had been given on a learning management system, but at present once the course finishes the student no longer has access. This was identified as a need to support ongoing student learning.

Educational aspects of feedback

Dialogue between students and teachers or their peers has been identified as a key component in providing effective feedback to students (Nicol & Macfarlane-Dick, 2006). A study by Orsmond, Merry, and Reiling (2005) noted that some students would like more opportunity to discuss feedback than they are given. It seems that e-tools may provide more opportunity for dialogue than traditional methods. O'Reilly (2005) noted various aspects of learning designs that become possible and manageable with electronic technologies. In support of this, a large number of lecturers described using online bulletin boards or forums to discuss assignments or questions about a course with the lecturer and other students. This provided students with the opportunity to ask questions of lecturers or tutors and have them answered online. It also gave students the opportunity to engage in discussion about the course or assignments with other students.

Nicol and Macfarlane-Dick (2006) emphasised attempting to close the gap between current and desired performance as a principle of good feedback. Although this was not an issue that was discussed by lecturers at length, some suggested that e-tools provide lecturers with the opportunity to compare past performance with current performance by virtue of the fact that they were able to keep an electronic archive of student’s past assignments. Having an archive facilitates accessing a student’s previous assignments and feedback that had already been provided to them. Some lecturers also discussed the importance of giving timely feedback so students can take into account the feedback they were given when working on a following assignment, which provides students with the opportunity to close the gap between their current and desired performance. This is a key part of learning and, as Orsmond et al. (2005) reported, a number of students read and reread comments. An example of a lecturer describing providing continuous feedback to students is presented below:

R5: My response is let’s have a number of assignments, and I think I had four, it might be five, that start off really simple. They build up over that period and there’s no point me giving back another assignment until the next one, unless the previous one is marked on there. So by sizing them and basically having some support to turn them round, you know I want to give them continuous feedback.

Reliability and validity in assessment procedures was another issue discussed by a large number of lecturers. The majority of lecturers discussed some form of keeping marking reliable, such as using marking schedules, or moderating marking before assignments were returned to students. This was particularly the case for larger classes, where tutors or markers were involved in marking. Some lecturers suggested that using e-tools for the moderation of assignments was particularly useful, as illustrated in the example below:

R2: Yeah I think it is easier for moderation purposes because like the same group we have internal classes as well as EDO [Electronic Delivery Option] classes. If electronically submitted then we can moderate each other’s work. Otherwise some are Word Documents and some are in hard copy and some are not and then there are problems in moderating also, because you can just send and it’s quick and faster and then we need not meet face to face….

Conclusions from interviews

The interviews provided evidence and real-world examples of e-tools being used in the assignment process. Survey participants recognise the benefits of technology in the marking and management of assignments. These range from
increased efficiency to more educationally effective approaches. Participants also identified benefits for the students such as saving time and reducing printing costs by electronically submitting, increased legibility of feedback comments, and the educational advantages of online discussion forums.

The survey identified a range of approaches to assessment. There is a core set of electronic tools such as email, word processors, spreadsheets and learning management systems through which participants are experiencing both efficiency and quality benefits. The extent to which more specialised tools for plagiarism, quality control, moderation, etc. are applied is limited, even amongst this enriched sample.

The decision to use e-tools has more to do with the individual than with faculty, class size, course level, or institution. There are clusters within some faculties where a culture of technology exists, where often an early-adopting or technology-minded person leads the uptake of new technology, sometimes even designing their own system. This eventually trickles down to the other staff members. Lecturers using technology often feel that there is a lack of support from their institution. This highlights the need for a support framework.

Lecturers wishing to integrate e-tools into the assignment process are faced with learning new technologies, acquiring new skills, and slightly altering their current practices. However, the learning curve is not too steep and, in the sample set, once the tools had been adopted, lecturers definitely did not want to revert back to a traditional paper-based approach.

**Recommendations**

An important goal of the research project underlying this article was to provide good practice for using e-tools for dealing with assignments. To meet this goal, practical recommendations building on the outcomes of literature research and survey were formulated. These recommendations were targeted at individual lecturers, allowing them to make changes in their assessment practices immediately and largely independently of additional institutional support.

**Mapping of tools to the stages of the assignment assessment process**

The assignment assessment process can be divided into seven stages. These stages are useful for understanding which way particular tools can support the process. The stages stem from the literature on key principles of good assessment (Nicol & MacFarlane-Dick, 2006) and the steps in dealing with assignments. The stages have been used and verified in an analysis of the interviews that examined current practice of the use of e-tools for assignment assessment (Milne, Heinrich & Morrison, 2008).

1. The process starts with supporting students with assignments such as clarifying what good performance is and allowing students to discuss the assignment goals and marking criteria so they understand them.
2. Assignments can be submitted using a variety of electronic tools.
3. The preparation of marking is about getting the assignments and organising them. Many of these tasks can be automatic or semiautomatic.
4. The marking stage is where the lecturers use their professional judgement to prepare feedback. There are a variety of tools to help the lecturer with marking but they all require the expertise of the human marker.
5. Keeping records includes recording and analysing marks and feedback, then transferring marks to other systems such as a central database.
6. Releasing results and feedback is when the lecturer provides the feedback and the marks to the students, facilitates discussion of the feedback, and provides opportunities for the student to incorporate the feedback into their work.
7. Using assignment experience for future teaching is the final phase. This is when the lecturer reflects on the process to provide further teaching for the students or to refine the activity for future student cohorts.

As outlined in the literature review section, the relevant e-tools can be divided into three groups:

- Generic office software and communication tools such as word processors, spreadsheets, and email
- Learning management systems such as Blackboard and Moodle
- Assignment assessment tools designed specifically to support lecturers with the marking and management of assignments.
Table 1 indicates how the tools of the first two groups can be used for the assignment management and marking processes. The specialist assignment assessment tools, which form the third group of tools, provide support at various stages. The following paragraphs indicate how several of these specialist tools link to the stages in the assignment assessment process outlined in Table 1.

Plagiarism-detection software such as Turnitin supports the preparation of marking by highlighting areas of suspected plagiarism. The system can also deal with the submission of assignments (Table 1, stages 2 and 3). Grademark supports the marking with inline comments and marking rubrics (but can only be used in conjunction with Turnitin, stages 4 and 6).

The electronic feedback software builds on Microsoft Office. It allows the preparation of standard and individualized comments, emails feedback to students, and analyses grade distributions. With these features the software addresses, the stages from marking to using the assignment experience for future teaching (stages 2, 3, 4, 6, and 7).

While Questionmark Perception is used mainly for automated assessment of multiple-choice and matching-type questions. It also has a tool for marking essay-type assignments. This tool is based on a marking rubric and allows comments within the student work as well as overall comments. It supports both marking teams and individual markers. The stages of preparation of marking and marking itself are addressed (stages 3 and 4) and, when used in conjunction with a learning management system, submission and release of results are covered as well (stages 1, 2, 5, and 6).

WebCTConnect works in conjunction with WebCT by extending the functionality of this learning management system. WebCTConnect prepares marking by recognising student groups and supporting marking teams. The marking is guided by marking schemes, and learning from the assignment experience for future teaching is addressed by extraction of all marking comments. WebCTConnect can be used in conjunction with MarkTool, which supports inline marking comments (WebCT: stages 1, 2, 5, and 6; WebCTConnect: stages 3 and 4; MarkTool: stage 4).

Markin starts by importing the student’s writing into the program. The marker is then able to mark the text using annotations, comments, and grades. The marked work is exported in a suitable format and sent to the student. When several pieces of work have been marked, one can use the program’s comprehensive statistics analyser to collate and compare marks from different students, or from different pieces of work by the same student. The program addresses the stages from marking to using the assignment experience for future teaching (stages 4, 5, 6, and 7).

<table>
<thead>
<tr>
<th>The stages of the assignment marking process</th>
<th>Generic office software</th>
<th>Learning management systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Supporting students with assignments</strong></td>
<td>Supports writing of assignment specification. Can email students about assignment.</td>
<td>Facilitates the delivery of assessment activities. Provides a common resource to explain goals and assessment rubrics. Integrates online discussions into course so students can discuss the assignment.</td>
</tr>
<tr>
<td>2. <strong>Submission of assignments</strong></td>
<td>Supports writing of assignment documents in generic formats. Students can email assignments.</td>
<td>Manages the electronic submission of assignments.</td>
</tr>
<tr>
<td>3. <strong>Preparation of marking</strong></td>
<td>Lecturer can email assignments to markers.</td>
<td>Allows groups to be set up. Assignments can be assigned to markers.</td>
</tr>
<tr>
<td>4. <strong>Marking</strong></td>
<td>Feedback and marks can be added to assignments.</td>
<td>Minimal support for document mark-up within the system.</td>
</tr>
<tr>
<td>5. <strong>Keeping records</strong></td>
<td>Lecturer can keep records using these tools with minimal support.</td>
<td>Extensive support for result aggregation storage. Often able to perform statistical profiling. Can be integrated with other ICT systems.</td>
</tr>
<tr>
<td>6. <strong>Releasing results and providing feedback</strong></td>
<td>Lecturer can email marked assignments to students.</td>
<td>Supports delivery of results generally through both email and web access.</td>
</tr>
<tr>
<td>7. <strong>Using assignment</strong></td>
<td>Minimal support for this stage.</td>
<td>Some provide rule-based tools for...</td>
</tr>
</tbody>
</table>
Markers Assistant is an application that focuses on the management of the assessment cycle. It facilitates the development of a marking rubric and the distribution of assignments to multiple markers. Once marked, the individual markers’ files are merged into a database. During marking, comments can be added to the submitted work. Post-marking moderation aids the quest for improved quality in marking. When results are returned to a student, only the comments that are associated with areas where the student has lost marks are provided to them (stages 3, 4, 5, and 6).

These specialist assignment assessment tools have considerably extended scope compared to generic software and learning management systems. Yet, there are several issues that limit the wider adoption of these tools. Most of the tools named require licence payments and/or are only available locally at the institutions where they have been developed. Besides payment, some level of technical and instructional support will be required by many lecturers to use these tools. Using a specialist tool necessitates an investment in time to learn the features of the tool.

Profiles, case studies, and action plan

The previous paragraphs have looked at the stages of the assessment process and at individual tools. The following paragraphs provide a different angle towards encouraging uptake of e-tool support. The idea is to build on the motivation, needs, and benefits as seen by the lecturers interviewed and to provide themes that individuals should be able to identify with. This leads to the formulation of three profiles that set out a pathway from relatively simple tool use to more sophisticated applications. Profile 1 focuses on efficiency improvements with the idea of freeing up time currently wasted on mundane administrative tasks. Profile 2 shifts towards quality improvements by suggesting how to use tools to contribute to better assignment assessment outcomes. These two first profiles only require the use of standard tools and features of widely available learning management systems. This ensures that individual lecturers should be able to get started without the need for additional licences, institutional support, and a steep learning curve. Profile 3 addresses the lecturer who is prepared to move beyond standard towards specialist tools and is ready to look at extensions of standard assignment assessment designs. Appendix 1 describes profile 3.

To complement the profiles, case studies and an action plan have been devised. The case studies describe typical teaching and assessment situations. Each case study depicts a scenario and discusses specific challenges and how to address these with the help of e-tools. The case studies have been derived from interview responses and are aligned with the profiles. An example of a case study relating to profile 3 is provided in Appendix 2. The action plan provides lecturers with a time-sorted checklist with the steps to take and when to take them. For example, assessment requirements, which concern students, have to be announced early on, sometimes even before the students enrol in a course. For electronic assignment submission, web access is needed, so this might be an enrolment consideration for students. Table 2 shows one section of the action plan. The full profile information, the case studies, and the full action plan are available from the project website (http://etools.massey.ac.nz).

Table 2. Section of the action plan for the step “marking of assignments”

<table>
<thead>
<tr>
<th>Tasks for step “marking of assignments”</th>
<th>Activities to commence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release of sample solution</td>
<td>You might want to release a sample solution to the class.</td>
</tr>
<tr>
<td>Preparing access to the assignments</td>
<td>If assignments are arriving by email, you will have stored them already in one dedicated place.</td>
</tr>
<tr>
<td></td>
<td>If you use the assignment tool of your LMS, you will need to download the assignments.</td>
</tr>
<tr>
<td></td>
<td>If you work with markers, you need to assign assignments to markers.</td>
</tr>
<tr>
<td>Marking</td>
<td>You use the tools and mechanisms selected for providing feedback.</td>
</tr>
<tr>
<td></td>
<td>If working with markers, you conduct your steps of monitoring markers and ensuring marking quality.</td>
</tr>
<tr>
<td></td>
<td>If you mark by yourself, you should guard against marker drift by revisiting assignments marked early on.</td>
</tr>
</tbody>
</table>
Overall evaluation  | Looking across all assignments, you can identify general strength and weaknesses and prepare material for further teaching accordingly. You might want to calculate some simple statistics for the assignments or look at the statistics the LMS has prepared for you.

Storing of marks  | You need to store all the marks for future reference and for import into your institution’s result processing system. Depending on the tools you have chosen, you might be able to simply export the marks in the required format.

Conclusion

This article has investigated assignment assessment that focuses on essay-type student work. The important characteristics of this type of student work are that students have to formulate their own thoughts and that the input of a human marker is required to understand this work and provide meaningful feedback, enabling student learning. The literature paints a clear picture of the importance of essay-type student work and the value of formative feedback for supporting learning.

A range of e-tools is available to support essay-type assignment assessment. Specialist tools exist that show high potential for effective support. The issue with these tools is that they are not widely accessible to individual lecturers and that they require a degree of institutional support. Together with the learning curve involved in mastering these tools, these are the reasons for a fairly low uptake despite the potential. Yet, even with tools that are part of widely available learning management systems and with general purpose tools like word processing and email, essay-type assignment assessment can be supported reasonably well. This arises from the literature and has been confirmed in interviews with lecturers across New Zealand tertiary education institutions. These lecturers, selected for their innovative use of e-tools, have described their use of e-tools in support of assignment assessment and have reflected on their marking and feedback practices. While these lecturers remember the challenges involved in the uptake of e-tools, they give the clear message that they do not want to revert to practices without e-tool use.

To facilitate the uptake of e-tools for assignment assessment, recommendations have been formulated. These recommendations address individual lecturers and show how these can use e-tools to support their existing assignment practices and work on improving their practices in keeping with educational theories. The recommendations take into consideration the fact that lecturers are pressured for time and do not necessarily have effective institutional support structures behind them. The recommendations target time savings that, in the next step of e-tool uptake, can contribute to improvements in quality. The initial steps only require access to widely available tools, lowering barriers to uptake based on institutional dependencies. It is hoped that these practical recommendations, which are based on solid academic research, will encourage many lecturers to increase their use of e-tools in assignment assessment, will contribute to an improvement in marking and feedback quality, and will ultimately lead to better student learning.

Acknowledgements

This project was funded by an e-Learning Collaborative Development Fund of the New Zealand Tertiary Education Commission. The contributions of Luanna Meyer and Bruce Granshaw (Victoria University of Wellington), Terry Crooks (University of Otago), and the research assistants Annabel Ramsay and David Morrison are acknowledged.

References


Appendix 1. Description of profile 3

Profile 3. Moving beyond standard tool use/extending assessment design: Stepping outside the conventional
The first two profiles describe the use of standard tools for a standard assignment process. The third profile will be of interest to you if you are prepared to step outside the conventional. This can be to explore specialist assignment tools or to continue with the standard tools but use them in more imaginative ways. Your motivation might be to provide further educational benefits, to gain efficiency, or to target both. As specialist assignment marking tools have been designed with marking problems in mind, a gain in both areas is possible.

Using non-standard tools will mean that you will have to gain access to these tools and be willing to experiment. These tools are not difficult to use but, as for any new tool, you will have to put some effort into exploration. It might pay to work on a smaller pilot first until you have gained confidence. Some of the tools will only affect you as the marker, and not the students. This will make it easier for you to experiment, because you won’t risk adverse reactions from the students.

You may have already thought about more involved assessment designs and quality control of the marking process, but might have hesitated because of the additional effort required. The use of standard tools beyond the conventional can help you to address these issues.

Most of the approaches discussed in this profile permeate all the stages of the assignment process outlined in Table 1. This section has therefore been structured by themes and not by stages.

Group assignments
Lecturers use group assignments for pedagogical or practical reasons. The questions that typically arise with group submissions are: Does only one group member submit the assignment on behalf of the group or does everyone submit? Are all submitted copies of one group identical? How do I ensure that all group members receive feedback? How do I make sure that I record marks for each group member? For large classes, these questions result in considerable administrative overheads.

Some of the specialist tools address these issues. They facilitate the grouping of students by either explicit definition of groups or by automated detection of group membership based on the assignment submitted. During marking, the tools associate feedback and marks with all group members, while giving you the option of providing the same or different marks to individuals. All group members will receive marking results automatically.

Working with markers
Working with markers poses challenges in administrative overhead and consistency control. Some of the specialist tools help you with these challenges. The tools facilitate the sharing of marking rubrics. They allow you to assign markers to assignments, for example through an initial random distribution which can be adjusted manually should the need arise. The marking of all the individual markers is integrated by the tools into one overview. This allows you to easily monitor progress and to implement schemes of consistency checking. As marking with electronic tools retains copies of all marked assignments you will have a record of who marked which assignments and of all marking comments. This is very helpful when students ask you for clarification, as you immediately have access to the marker’s comments. Combining all markers’ feedback and marks, in total and per marking rubric, into one environment assists with the analysis of the overall class performance. This in turn helps in your teaching preparation for current and future classes.

Plagiarism
Unfortunately plagiarism can be a major issue, with students copying work among each other and from external sources. Plagiarism detection tools help you to spot potential plagiarism. The tools compare student work on a class level and against ever increasing repositories of previous student work and publicly available material. Some of the plagiarism detection tools are combined with marking tools, addressing questions of marking rubrics, feedback and bookkeeping. You can also use a combination of tools, a specialist tool for plagiarism detection followed by the marking tool of your choice.

Focus on the formative process
One common concern regarding assignment marking is that students do not pay attention to the feedback provided to
them and therefore the formative effect is lost. To combat this problem, one can extend the standard assignment process of submission and return to encourage involvement with the feedback. This can mean returning marked assignments with only formative feedback initially, with marks either delayed or provided only after a student responds to the feedback. Some lecturers allow resubmission of assignments in response to formative feedback. Peer-assessment processes can encourage students to engage with the assignment topics on a deeper level.

While some of the specialist tools support such processes directly, you can already achieve a lot using standard tools like LMS. For example, you could set up a two-part assignment, where the first submission serves largely formative purposes and the second serves as the revised version. In a similar fashion, the first submission could tackle a standard assignment, while the second submission could focus on students’ assessment of the work of their peers. Using a standard LMS you would have to perform the intermittent step of making confidential assignment submissions available to the whole class, for example via the LMS content modules. Such processes could be complemented by the use of the LMS communication tools. In selecting which tools to apply and how to combine their use it is important to consider their individual strengths, such as confidentiality provided by an assignment tool, open access at the class level for a content module, or shared contributions in a discussion environment.

**Ways of providing comments**

What has been discussed in the previous profiles is to provide written comments either by writing directly on the student’s work or by attaching separate marking sheets. Some specialist marking tools add to this by allowing the marker to link their comments to the categories of the marking rubric. The advantage for the student is that the marking comments are related both directly to the context in the student’s work and to the marking criteria. Some tools offer special pre- or marker-defined symbols for spelling or grammatical mistakes, for example. Another way in which some tools assist the marker is by providing various graphical highlights to clearly identify the reference point in the student’s work.

Audio comments can be an alternative to written feedback. Through standard word-processing software, audio comments can be inserted much like written comments at specific reference points in the student’s work. This might be an alternative for markers who have problems with typing a large number of comments, for situations where audio comments suit the subject area, or simply to create student interest in the feedback by diverting from the norm of written feedback.

**Further efficiency gains**

Assignment submission and management with standard LMS tools is already more efficient than email and paper submission. Still, specialist tools provide further advantages. For the submission of multiple files, students are often asked to combine these files into a single compressed file. While standard LMS can accept such files, they do not automatically uncompress back to multiple files. This means manual and repetitive work for the marker who is not supported by a specialist tool that can perform these tasks automatically.

It is helpful to record the marking status of an assignment as part of the marking process. In the context of quality assurance it can be required to revisit an assignment marked earlier on or to review marking done by markers. Some specialist tools allow the marker to capture such meta-information within the marking environment. This integrated approach increases the efficiency and in turn makes it more likely that quality assurance steps are actually undertaken.

**Further quality gains**

Opportunities for quality gains in terms of coordinating and monitoring the work of markers and keeping information of the marking status of assignments have already been mentioned. What can be added here is the use of standard LMS features to involve students in the development of the assignment task specification and the marking rubric. Students can use discussion tools to exchange ideas and comments. The lecturer can use the content modules of the LMS to make draft documents available to students, for example after a class discussion on assignment goals and student work towards a possible marking rubric.
Appendix 2. Case study relating to profile 3: Providing audio comments

Introduction
I am a communications lecturer teaching generic communications courses across a range of subjects. My students come from a wide variety of backgrounds, some being technically oriented whilst others being artistic. I also have a number of students for whom English is a second language. This variety of styles and cultures gives me many challenges when I’m providing constructive feedback.

The need to change
I started using audio feedback for students years ago, using a simple cassette tape recorder. I was motivated to use audio because I felt so limited by the lack of space for adding comments to assignments and also the time it took to provide written feedback. In communications it’s common for each assignment to go through a number of iterations, drafts, I guess you could call them, with each reviewed and commented on. Trying to explain the real issues using text-based feedback often took a lot of effort because I had to make sure the emphasis of the comment was unambiguous and constructive. I find using audio feedback much faster and more flexible as it allows me to provide the sort of dynamic feedback that helps my students.

Marking using audio feedback
My colleagues and I use audio formative feedback for all our students regardless of discipline. So the technical trade students get it as do the business students. Typically I give 10 minutes audio feedback per assignment and this lets me give much richer feedback than I could when I was limited to text only. I’m able to offer them alternate ways of expressing a point and demonstrate how a particular statement could be ambiguous and how small changes can make a big difference. I also find, and the students tell me they do too, that audio feedback feels a lot less formal, a more conversational way of communicating feedback. Although I haven’t done any formal evaluation, anecdotally I do think I get fewer complaints or queries over the feedback I give and I am convinced that I spend much less time marking.

The process
When I mark an assignment I start by saying the student name and description of the assignment. Depending on the tool I’m using, I either pre-save the audio file named with the student identification and date or name it when I’m finished. That’s because some standalone recorders don’t allow you to name the file at the outset. I then read through the assignment and when I get to a place that I want to comment on I start the recorder and add the comment. To help the student know what part of their work I am commenting on, I use one of the following aids. If the section is easily identifiable (e.g., close to a paragraph heading) I just quote the part of the text I’m going to comment on and then add my comments. More frequently I mark the script with a number or a letter and then refer to that in my commentary. Once the audio is complete I save it and then either email it to the student with a marked-up version of their work or burn it onto a CD and mail it to them. Mostly I use the mp3 format. I have heard that you can add audio comments to Word documents but I haven’t tried that myself. Worth a look though.

The technology
I’ve tried quite a few technologies for recording feedback and most work just fine. Even the cassette tape worked but the tapes were quite expensive. Thanks to the popularity of the mp3 format you can get all sorts of tools, many for free. I’ve used Garage Band on my Apple laptop and one of my colleagues uses Audacity on his PC. Audacity is a free recording application that is also really easy to use. You need a microphone, of course, and I find a headset really useful, especially in a shared office. I have also used standalone audio recorders, modern dictation machines like the Olympus DS. There seems to be a solution for most budgets.

Challenges with using audio feedback
Because the audio feedback I use is separate from the original document I do need to be careful about the labeling of the audio files. When I first started, I did find that I needed to practice getting an appropriate narrative style but that doesn’t take long to master. I use two methods depending on which one suits the situation. Sometimes I read the assignment right through first and make hand written notes but that method has its drawbacks. Not only does it take more time, there is a tendency to read your notes and this removes a lot of the richness you get from doing it live! I think it helps if you are a confident and clear speaker, not too many “you knows” or “umms” but I think that’s something that develops as you get more practice. In fact it can actually help you improve your own communication skills!
One technical challenge of using audio feedback is the size of the file that is generated. Using the mp3 file format the file size is often about 1 MB per minute of recording, although this can be reduced by changing the export file’s settings. In Garage Band, exporting an audio file as a mono podcast will give about 250 KB/minute. If you want to know more I suggest you do a web search on “reducing audio file size” or similar key words because there is a lot of advice available.

Evaluation
When we work with a class for the first time, we do have an evaluation questionnaire we ask the students to fill in. Audio feedback doesn’t suit everyone, but in the 2 or 3 cases where students have really objected, we provide them with written feedback. In the end-of-course evaluation, our department has added a section that deals with just the audio feedback and how it can be improved. That way we can keep track of the shifts in the way our students use technology and we can adapt what we do.

Summary
By using audio feedback I believe I can provide much richer and more useful comments and suggestions for my students. I am certain that if you were to compare like with like, audio feedback takes much less time than written feedback. It is important that you adopt a style that works for both you and your students and I do modify my style for each student group. The cost is minimal for both consumables and the equipment itself so that hasn’t been a barrier for us. With the number of iPods and phone mp3 players around now, students have told me that audio feedback is making their learning experience much more flexible too.
Multi-Objective Parallel Test-Sheet Composition Using Enhanced Particle Swarm Optimization

Tsu-Feng Ho\textsuperscript{1}, Peng-Yeng Yin\textsuperscript{2}, Gwo-Jen Hwang\textsuperscript{*3}, Shyong Jian Shyu\textsuperscript{1} and Ya-Nan Yean\textsuperscript{2}

\textsuperscript{1}Department of Computer Science and Information Engineering, Ming Chuan University, Taoyuan 33348, Taiwan
\textsuperscript{2}Department of Information Management, National Chi Nan University, Nan-Tou 54561, Taiwan
\textsuperscript{3}Department of Information and Learning Technology, National University of Tainan, Tainan City 70005, Taiwan

\texttt{ttho@mail.mcu.edu.tw // ppyin@ncnu.edu.tw // gjhwang@mail.nutn.edu.tw // sjshyu@gmail.com

mandywang@xic.com.tw}

\textsuperscript{*}Corresponding author

ABSTRACT

For large-scale tests, such as certification tests or entrance examinations, the composed test sheets must meet multiple assessment criteria. Furthermore, to fairly compare the knowledge levels of the persons who receive tests at different times owing to the insufficiency of available examination halls or the occurrence of certain unexpected situations, a set of parallel test sheets needs to be composed, which is almost impossible to accomplish manually. To cope with this problem, an enhanced particle swarm optimization approach is proposed to efficiently compose parallel test sheets from very large item banks, while simultaneously meeting multiple assessment criteria. Moreover, a computer-assisted testing system was developed, and a series of experiments were performed to evaluate the comparative performance of the proposed approach and a genetic algorithm counterpart.

Keywords
Test-sheet composition, Computer-assisted testing, Particle swarm optimization, Genetic algorithm

Introduction

The key to a good test depends not only on the subjective appropriateness of test items, but also on the way the test sheet is composed. As the number of test items in an item bank is usually large, and the number of feasible combinations to form test sheets grows exponentially, it thus takes an extremely long time to set up an optimal test sheet according to the given objectives. For a formal and large-scale test, such as national certification tests or entrance examinations, the test-sheet composition problem is even more difficult to cope with, since a set of parallel test sheets needs to be composed for handling any kind of accident; for example, if some areas of the country cannot hold the test owing to inclement weather, alternative test sheets with different test item compositions but identical test ability are needed for later tests.

Previous investigations have shown that a near-optimal solution for composing one test sheet is difficult to find when the number of candidate test items is larger than five thousand (Hwang, Lin, Tseng, & Lin, 2005); thus, it is much more difficult to compose a set of near-optimal parallel test sheets from a very large item bank. In this paper, we describe this problem in a general and formal way by which multiple objectives and a set of criteria can be defined by the examination committee. For instance, the test abilities of these parallel test sheets are set as close as possible while satisfying the constraints entailed by the expected testing time, the degree of discrimination, and the expected degree of relevance to each specified concept. To cope with this complex problem, an enhanced multi-objective particle swarm optimization (EMOPSO) algorithm is proposed. By employing this novel approach, the generated composition of parallel test sheets can optimize multiple objectives and meet the needs of a specified set of criteria.

Background and related works

Test-sheet composition

Research indicates that a well-constructed test sheet not only helps in evaluating the learning achievements of students, but also facilitates the diagnosis of the problems embedded in the learning process (Hwang, 2003a, 2003b; Hwang, Tseng, & Hwang, 2008). Moreover, selecting proper test items is critical in creating a test sheet that meets multiple assessment criteria, such as the expected time needed for answering questions on the test sheet, the number
of test items, the specified distribution of course concepts to be learned, and the expected discrimination or degree of
difficulty (Hwang et al., 2005; Hwang, Lin, & Lin, 2006; Yin, Chang, Hwang, Hwang, & Chan, 2006).

Most of the existing approaches for test-item selection focus on the composition of a single test sheet. However, for
formal and large-scale tests, it is necessary to compose parallel test sheets with different sets of selected items such
that the test ability (in terms of appropriate measures) of each sheet is as similar as possible. It is obvious that the
problem of parallel test-sheet composition is much more difficult than that of single test-sheet composition;
therefore, a more efficient and effective approach is needed for the former case.

Pareto optimal solutions

Single-objective optimization problems (SOOP) seek to optimize a single goal, such as cost minimization. However,
multiple-objective optimization problems (MOOP), which simultaneously optimize a set of goals, are more suited to
describing real-world applications. Therefore, MOOP is more computationally intensive than SOOP. A notion in
MOOP is to provide the decision-maker with a set of Pareto optimal solutions, which are not dominated by any other
solutions. Solution \( x \) is said to dominate another solution \( y \) if \( x \) is strictly better than \( y \) based on the objective value
for at least one objective and \( x \) is no worse than \( y \) for the others. Obviously, if the final solution is not Pareto optimal,
它可以改进至少一个目标而不危及其他目标的质量。

Most existing methods for tackling MOOP apply the weighting sum method (Lai & Li, 1999), i.e., transforming the
multiple objective functions into a single objective function by calculating their weighting sum. However, multiple
runs of the weighting sum method are needed to obtain a set of non-dominated solutions. This drawback is
experienced by several other techniques such as goal programming (Schniederjans, 1995) and epsilon programming
(Yokoyama, 1996). Recently, metaheuristic-based algorithms have been introduced as a viable technique to tackle
MOOP due to its single-run advantage.

Particle swarm optimization

Particle swarm optimization (PSO) is biologically inspired, drawing on the observation that birds flock
synchronously, change direction suddenly, scatter, and regroup again (Clerc & Kennedy, 2002). Each individual
entity, called a particle, benefits from its own experience and that of the other members of the flock during the
foraging process in order to increase the chance of finding food.

<table>
<thead>
<tr>
<th>1. Initialization.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Generate ( P ) particle positions, ( \hat{P}<em>i = (p</em>{i1}, p_{i2}, ..., p_{ib}) ), ( 1 \leq i \leq P ), at random.</td>
</tr>
<tr>
<td>1.2 Generate ( P ) velocity vectors, ( \hat{V}<em>i = (v</em>{i1}, v_{i2}, ..., v_{ib}) ), ( 1 \leq i \leq P ), at random.</td>
</tr>
<tr>
<td>2. Iteration (repeat until a stopping criterion is met).</td>
</tr>
<tr>
<td>2.1 Evaluate the fitness of each particle ( \hat{P}_i ).</td>
</tr>
<tr>
<td>2.2 Determine the best vector ( p_{\text{best}} ), visited so far by each particle ( \hat{P}_i ).</td>
</tr>
<tr>
<td>2.3 Determine the best vector ( g_{\text{best}} ) observed so far by the entire swarm.</td>
</tr>
<tr>
<td>2.4 For each particle ( \hat{P}_i ), update its velocity vector</td>
</tr>
<tr>
<td>2.5 For each particle ( \hat{P}_i ), update its particle position</td>
</tr>
</tbody>
</table>

Figure 1. Summary of the Type 1“ constriction PSO algorithm

Many variants have been proposed since the first PSO proposal; however, most of them resemble the Type 1“
constriction model (Clerc & Kennedy, 2002), which is summarized in Figure 1. Consider an optimization problem
with \( b \) decision variables. The algorithm initiates a swarm of \( P \) particles (candidate solutions) generated at random,
\( \hat{P}_i = (p_{i1}, p_{i2}, ..., p_{ib}) \), \( 1 \leq i \leq P \). A velocity vector \( \hat{V}_i = (v_{i1}, v_{i2}, ..., v_{ib}) \) is randomly created for each particle. In each
iteration, the fitness of each particle is evaluated. Then, the particle’s personal best \( (p_{\text{best}}) \) and the swarm’s best
\( (g_{\text{best}}) \) are identified. To construct the search course, for each particle we update its velocity \( \hat{V}_i \) and position \( \hat{P}_i \)
through each variable dimension \( j \) using Equations 1 through 3, as follows,
1. \[ v_{ij} \leftarrow K(v_{ij}+\varphi_1 \text{rand}_1[p_{bestj}-p_{ij}]+\varphi_2 \text{rand}_2[g_{bestj}-p_{ij}]) \]

2. \[ K = \frac{2}{\varphi_1 + \varphi_2 - \sqrt{(\varphi_1 + \varphi_2)^2 - 4(\varphi_1 \varphi_2)}}\]

3. \[ p_{ij} \leftarrow p_{ij} + v_{ij} \]

where \( \varphi_1 \) and \( \varphi_2 \) are the cognitive coefficients, \( \text{rand}_1 \) and \( \text{rand}_2 \) are random real numbers having values between 0 and 1, and \( K \) is the constriction coefficient. In essence, the particle explores a potential region defined by \( p_{best} \) and \( g_{best} \), while the cognitive coefficients and the random multipliers change the weightings for the two bests in each iteration. The use of the constriction coefficient is to ensure the convergence of the algorithm. Typically, \( \varphi_1 + \varphi_2 \) is set to 4.1 and \( K \) is thus 0.729. The PSO algorithm must be terminated with a maximum number of iterations or the overall best particle cannot be further improved.

PSO has exhibited great success in solving SOOP, ranging from evolving weights and structures for artificial neural networks (Eberhart & Shi, 1998), to image registration (Yin, 2006), to task allocation (Yin, Yu, Wang, & Wang, 2007). Nevertheless, there are very few PSO variants devoted to MOOP. Coello Coello, Pulido, & Lechuga (2004) proposed a multi-objective PSO (MOPSO) algorithm that maintains an external repository that stores the non-dominated solutions observed overall, and randomly chooses one member from this repository to serve in place of the swarm’s best solution. Mostaghim, Halter, & Wille (2006) proposed a linear multi-objective particle swarm optimization (LMOPSO) to deal with optimization problems with linear constraints. The LMOPSO computes the basic feasible solutions using linear programming and finds a random set of feasible solutions (particles) based on linear combinations of basic feasible solutions so that the produced solutions never become infeasible. However, the real-world, multi-objective parallel test-sheet composition (MOPTSC) problem we consider in this paper is not suited to being solved by the LMOPSO for the following reasons:

a. The MOPTSC problem involves a tremendous number of decision variables and constraints. The computation resources consumed by linear programming would already be more expensive than those needed by our EMOPSO algorithm.

b. Our problem formulation considers integer-value variables instead of real-value variables, only the integer vectors within the polyhedron formed by the linear constraints may correspond to feasible solutions (some integer vectors are forbidden by constraints 9 and 10 in our problem formulation, as will be discussed in the next section.) In other words, the linear combinations of basic feasible solutions very likely produce infeasible solutions, and a post-repairing algorithm is still needed to preserve the feasibility of the produced solutions.

**Problem description**

**Multi-objective formulations**

We consider the multi-objective parallel test sheet composition (MOPTSC) problem that requests a test design with identical test abilities (discrimination and degrees of difficulty) while respecting multiple constraints (assessment time and the degree of relevance to each specified concept). Given a large test bank consisting of \( N \) items, namely \( Q_1, Q_2, \ldots, Q_N \), we want to compose \( K \) parallel test sheets simultaneously. Each of the test sheets contains exactly \( S \) test items selected from the test bank, such that for any pair of test sheets \( k \) and \( l \), \( 1 \leq k \neq l \leq K \), there is no common test item. Each test item, say, \( Q_j \), is characterized by a set of attributes consisting of the expected time \( t_i \) for students to answer \( Q_j \), the degree of difficulty \( f_i \), the degree of discrimination \( d_i \), and the degree of association \( r_{ij} \) between \( Q_j \) and concept \( C_j \) (that is, \( r_{ij} \) denotes the relevance of \( Q_j \) to \( C_j \)). When composing each test sheet, a number of \( M \) course-related concepts, namely, \( C_j \), \( 1 \leq j \leq M \), should be covered by respective lower bounds \( h_j \), \( 1 \leq j \leq M \), and the expected time needed for answering each test sheet should be within the range of \( l \) and \( u \). In the following, a multi-objective, mixed-integer programming model is proposed to formulate the MOPTSC problem.
The decision variable $x_{ik}$ is 1 if test item $Q_i$ is included in test sheet $k$, and 0 otherwise. The first objective aims to minimize the maximal difference of the average discrimination degree between any two parallel test sheets, and the second objective seeks to keep the maximal difference between the average difficulty degree of the test sheets and a given target difficulty $F$ as small as possible. The two objectives are designed with regard to the considerations we made in the real scenario of preparing national test sheets.

The problem constraints are explained as follows. Constraint 6 stipulates that the concept relevance degree of the selected test items on each generated test sheet must sum up to a value no less than the expected degree of relevance to each specified concept. Constraints 7 and 8 indicate that the expected test time of each generated test sheet must be in its specified time range. Constraint 9 requests that no pair of test sheets contain any common test item. Finally, Constraint 10 states that all of the test sheets must contain the same given number of test items.
Enhanced multi-objective PSO method

Enhanced multi-objective PSO for the MOPTSC problem

This study proposes the EMOPSO (enhanced multi-objective PSO) algorithm for finding quality solutions to the MOPTSC problem. The pseudo-code of EMOPSO is outlined in Figure 2.

Algorithm: EMOPSO

**Input:** \(N\) test items \(Q_1, Q_2, \ldots, Q_N\) and \(M\) course concepts \(C_1, C_2, \ldots, C_M\);  **Output:** \(K\) parallel test-sheet composition.

**Step 1. Initialization**

The candidate solution to the MOPTSC problem consists of \(K\) test sheets, each of which selects exactly \(S\) test items from the test bank. Thus we represent a particle (a candidate solution) by a vector of \(K \times S\) elements where each element can take a value ranging from 1 to \(N\) (size of the test bank) while respecting constraint 9 that no chosen item appears more than once in the given particle. Formally, the \(i\)th particle is represented by

\[
P_{\rho i} = [p_{i1}, p_{i2}, \ldots, p_{iS}, p_{i1}, p_{i2}, \ldots, p_{iS}]\]

where \(p_{kj} \in [1, N]\) indicates the original index of the chosen item in the item bank to compose the \(s\)th item for the \(k\)th test sheet. Using the terminology \(x_{ik}\) defined in our problem formulation, it is clear that \(x_{ik} = 1\) if \(i = p_{kj}\) for \(\forall k, \forall s\), and \(x_{ik} = 0\), otherwise. To prepare the initial particle swarm, \(P\) particles are generated at random. The initial \(P\) velocities are also generated at random.

**Step 2. Iteration**

**Step 2.1 Evaluate the multiple fitness values of each particle**

Our MOPTSC problem is composed of multiple objectives and a set of difficult constraints; thus, it is natural to consider multiple fitness functions to measure the particle quality. The first two fitness functions correspond to the two objectives of the MOPTSC problem:

11. \(Z_1(\vec{p}) = \max_{i=1}^{N} \left\{ \frac{\sum_{i=1}^{N} d_i x_{ik} - \sum_{i=1}^{N} d_i x_{li}}{\sum_{i=1}^{N} x_{ik} - \sum_{i=1}^{N} x_{li}} \right\} \)

and

12. \(Z_2(\vec{p}) = \max_{i=1}^{N} \left\{ \frac{\sum_{i=1}^{N} f_i x_{ik} - F}{\sum_{i=1}^{N} x_{ik}} \right\} \)

To account for the particle quality with respect to the ability to satisfy the problem constraints, we define the amount of constraint violation as the third fitness of the particle. As previously noted, it is very likely that infeasible solutions are produced. Transforming the constraints into another problem objective is beneficial to move the particles towards feasible regions. As the satisfaction of constraints 9 and 10 is assured by the particle representation scheme (see Step 1), we only need to deal with each of constraints 6–8 as follows:

13. \(R(x) = \sum_{k=1}^{K} \sum_{j=1}^{M} \max \left\{ 0, \left( h_{j} - \sum_{i=1}^{N} t_{ij} x_{ik} \right) \right\} \)

14. \(L(x) = \sum_{k=1}^{K} \max \left\{ 0, \left( t_{ik} - \sum_{i=1}^{N} x_{ik} \right) \right\} \)

15. \(U(x) = \sum_{k=1}^{K} \max \left\{ 0, \left( \sum_{i=1}^{N} x_{ik} - u \right) \right\} \)

Then we define the third fitness function of the particle as the weighting sum of these three violation amounts,

16. \(Z_3(\vec{p}) = w_1 R(x) + w_2 L(x) + w_3 U(x)\)

where \(w_1, w_2,\) and \(w_3\) denote the normalization factors for the three amounts in order to scale them onto the same value range \([0, 1]\) so that the third fitness will not be dominated by either of the constraint violations. Note that the
third fitness indicates the feasibility of the solution decoded from the particle. The resulting solution is infeasible if
\[ Z_3(\hat{P}_i) > 0, \]
whereas it is feasible only if \( Z_3(\hat{P}_i) = 0 \).

There exist several constraint-handling strategies in the population-based approaches like penalty functions, domination tests, stochastic ranking, etc. They should be modified to be applied in the MOOP context. For example, the original stochastic ranking proposal (Runarsson & Yao, 2000) with mono-objective scenario compares two infeasible solutions by associating a probability \( P_f \) to the case of using only the single objective and \( 1 - P_f \) to the case of using the overall amount of constraint violation. Considering the stochastic ranking with a multi-objective scenario, we can associate a separate probability to each of the objectives and still use a single objective for comparison. Here, we apply an advanced constraint-handling strategy originally introduced by Deb (2001) that treats the weighting constraint as another objective, and the solutions are compared based on the dominance relationship.

To be specific, there are three scenarios in which we compare two candidate solutions, say, A and B, as follows:

**Scenario 1:** Both A and B are feasible solutions.
Rule: Compare A and B based on the dominance relations using fitness functions \( Z_1(\hat{P}_i) \) and \( Z_2(\hat{P}_i) \) and retain the non-dominated solution(s).

**Scenario 2:** A is a feasible solution and B is an infeasible solution.
Rule: To rule out the infeasible solutions during the particle swarm evolution, we assume that a feasible solution A always dominates an infeasible solution B.

**Scenario 3:** Both A and B are infeasible solutions.
Rule: Compare A and B based on the dominance relations using fitness functions \( Z_1(\hat{P}_i), Z_2(\hat{P}_i), \) and \( Z_3(\hat{P}_i) \), and retain the non-dominated solution(s).

**Step 2.2 Update the elite set \( E \) according to the dominance relationship and crowding distance.**
In order to preserve the diversity of Pareto optimal solutions, we dynamically maintain an elite set \( E \) that selects the non-dominated solutions using the crowding distance technique. Initially, \( E = \emptyset \). Once a new trial solution is produced, it is compared to the incumbent members of \( E \) for possible inclusion in this set based on the dominance comparison. Note that some existing members of \( E \) have to be removed due to the inclusion of a new member if the former members are dominated by the new solution. We apply the crowding distance technique (Deb, Prarap, Agarwal, & Meyarivan, 2002) to estimate the density in the objective space surrounding each efficient point obtained by the non-dominated solution. The crowding distance of the efficient point is defined as the average distance of two points on either side of this efficient point along each of the objectives. When the elite set gets full, the non-dominated solution with the efficient point having the shortest crowding distance is removed, so the diversity of the members in \( E \) is preserved.

**Step 2.3 If \( p_{best} \) of particle \( \hat{P}_i \) has not been improved for \( \eta \) consecutive iterations, restart this particle using an elite solution.**
To explore the uncharted regions in the solution space (a diversification strategy), the search trajectory of a particle is reinitiated if this particle fails to improve its best solution for the past \( \eta \) consecutive iterations. When a particle is determined to be restarted, a member uniformly selected from the elite set is used as the seed for generating this particle and its best solution. The restarting strategy avoids the stagnation when the search gets stuck in the local optimum.

**Step 2.4 Use the Diversity strategy to choose a non-dominated solution visited so far by particle \( \hat{P}_i \) to serve as this particle’s best solution, \( p_{best} \).**
For the single objective PSO, the determination of \( p_{best} \) is simply according to the unique fitness value. However, for the enhanced multi-objective PSO involving three fitness functions, we have to determine \( p_{best} \) according to the dominance relations. In general, there may be more than one non-dominated solution identified in the search trajectory of a particle, so we store all of them in an archive. To guide the search of individual particles to different regions in the search space, which would enhance the diversity of the finally obtained solution front, we apply the Diversity strategy proposed in Branke and Mostaghim (2006). The Diversity strategy stipulates that each particle
selects from its personal archive the non-dominated solution that is the farthest away from the other particles in the objective space as the personal best for this particle.

Step 2.5 Use the Sigma strategy to choose a member elite from $E$ to serve as the swarm’s best vector, $gbest$.
Analogous to the determination of $pbest_i$, there is in general more than one non-dominated solution observed overall by the entire swarm. We therefore choose an elite solution from $E$ to serve as $gbest$ using the Sigma strategy (Mostaghim & Teich, 2003). For a focal particle, the Sigma strategy selects a non-dominated solution from $E$, which is the closest to the line connecting this particle’s position to the origin in the objective space. Therefore, the Sigma strategy lets the particles fly directly toward the Pareto-optimal front, and thus the convergence of the solutions is improved.

Step 2.6 Update the particle’s velocity and position.
After determining the best solution guides, $pbest_i$ and elite (serving as $gbest$), the particle’s velocity and position are updated using equations 1–3. In this way, the proposed EMOPSO algorithm is able to handle multi-objective problems without having to design new trajectory updating formulas. It is only necessary to determine appropriate solutions to play the roles of guiding solutions ($pbest_i$ and $gbest$).

Step 2.7 Mutate the particle.
Fieldsend and Singh (2002) introduced a term, turbulence, which disturbs the flight of particles by additionally adding a stochastic variable to the velocity obtained after applying the update. They empirically showed that the use of turbulence can be a significant aid to multi-objective PSO in solving several test functions. Coello Coello et al. (2004) have also empirically shown that additionally applying a genetic mutation operation in their MOPSO algorithm is helpful in finding a diverse set of Pareto optimal solutions. As a consequence, we implement a mutation operator that randomly chooses some variables of a given particle and replaces their values with alternative ones. The mutation operation is performed with a mutation probability that decreases with the number of evolutionary iterations.

Competing multi-objective genetic algorithm
A modified genetic algorithm (GA) is used to validate the feasibility of our EMOPSO algorithm in solving the MOPTSC problem (Hwang, Chu, Yin, & Lin, 2008). To make a fair comparison, the GA-based method employs the same population size, solution representation, elite set update with crowding distance control, and dominance comparison strategy as those used by the EMOPSO algorithm. The rest of the parameters and strategies follow the common practice used in the GA community. In particular, the top 10 percent elitist chromosomes from the current population are retained in the next population in order to ensure that good solutions will not be dropped due to randomization. The other 90 percent of the next population are determined using the tournament selection scheme. Uniform crossover is adopted where the allele for each gene of the offspring has an equal probability of retaining the value from either of its parental chromosomes. Then the mutation operator, which occasionally changes the allele to any feasible value, is performed. It is noted that the tournament selection scheme is conducted by reference to the GPSIFF fitness (Yen, 2003), which has been shown to perform comparably well as the non-dominated sorting employed by the Nondominated Sorting Genetic Algorithm II (NSGA II). For a given chromosome $\hat{p}$, its GPSIFF fitness is defined as $\text{GPSIFF}(\hat{p}) = p - q + D$, where $p$ and $q$ are the numbers of chromosomes in the current population that are dominated by $\hat{p}$, and those that dominate $\hat{p}$, respectively, and $D$ is a constant stipulating that the GPSIFF fitness never has a negative value.

System administration and comparative performance
The proposed algorithm has been implemented in a web-based learning diagnosis system. The interface of the system and the experiments for evaluating the comparative performance of the proposed approach are presented in the following subsections.
Web-based system administration

The web-based learning diagnosis system contains several online tutor assistant functions, including item bank management, concept management, test-sheet composition, conducting group tests, and analyzing learning status. To execute the EMOPSO algorithm, the composition criteria (equations 6–10) for the parallel test sheets should be specified through the interface depicted in Figure 3.

After giving instances of the composition criteria, the user can activate our EMOPSO algorithm to find near-optimal parallel test-sheet compositions. The composed test sheets can be browsed via our web-based interface, as shown in Figure 4.

Figure 3. Test-sheet composition interface for specifying part of the composition criteria

Figure 4. Interface for browsing the test sheets composed by the EMOPSO algorithm
Performance evaluation

Following are the performance metrics used in the experiments, the comparative results obtained by the competing algorithms on several item banks of variable sizes, and the sensitivity analysis results:

Performance metrics

There are two issues related to the performance evaluations for multi-objective algorithms: (1) convergence is concerned with how close the derived solutions are to the real Pareto optimal solutions, while (2) diversity refers to how well the derived solutions are separated. For these purposes, three performance metrics that are widely used in the literature (Coello Coello et al., 2004) are adopted and defined as follows:

Generational distance (GD). GD estimates how far the solutions generated by a given algorithm are from the nearest Pareto optimal solutions. Assuming that there are \( n \) solutions generated, for each of the solutions, let \( \Delta_i \) be the distance between this solution and its closest Pareto optimal solution. GD is calculated by

\[
GD = \frac{\sum_{i=1}^{n} \Delta_i}{n}.
\]

The smaller the value, the better the generated solutions.

Spacing (SP). SP measures the spread of the non-dominated solutions generated by a given algorithm. Let \( n \) be the number of generated solutions, \( d_i \) the distance between one of these solutions and its closest member of the generated solutions, and \( \bar{d} \) the mean of all \( d_i \). SP is defined as

\[
SP = \sqrt{\frac{1}{n-1} \sum (d_i - \bar{d})^2}.
\]

A set of generated solutions with a smaller SP value is desired because these solutions exhibit a better representation of a front. The minimum value of SP is zero, which indicates that all the generated solutions are equidistantly spaced.

Error ratio (ER). ER corresponds to the missing ratio that accounts for the number of generated solutions being non-Pareto optimal, divided by the total number of generated solutions. Again, let \( n \) be the number of solutions produced by a given algorithm, and \( e_i \), a binary variable which is equal to zero if the \( i \)th generated solution is a member of the true Pareto optimal set; otherwise, it is equal to one. ER can be derived using

\[
ER = \frac{\sum e_i}{n}.
\]

The smaller the ER value, the more likely the generated solutions are to be Pareto optimal.

Comparative results

In addition to the EMOPSO and the GA-based method, we further implemented a random search method, denoted by Random, which uniformly samples solutions in the search space. We executed the random search program by consuming the same number of function evaluations as the other competing algorithms. To test the robustness and effectiveness of the competing algorithms, 10 test item banks of variable sizes were produced. An item bank size of 250 to 50,000 items satisfies the needs of a wide range of real-world examination applications. We begin by describing the experimental results for composing three parallel test sheets, each consisting of 30 test items. The sensitivity analysis against other varying experimental conditions is noted in sensitivity analysis.

The three performance metrics (GD, SP, and ER) are based on the assumption that the set of true Pareto optimal solutions to the underlying problem is known. Nevertheless, for our MOPTSC problem involving extremely high dimensions and complex constraints, it is almost impossible to list all the true Pareto optimal solutions. Alternatively, a competitive evaluation strategy is introduced to remedy this problem. Each of the competing algorithms was executed for multiple independent runs. The pool of all the derived solutions from these runs was examined, and the non-dominated solutions by reference to this pool were considered approximate Pareto optimal solutions, which were then used to evaluate the performance metrics of each separate algorithm. In our experiments, five multiple runs of each of the EMOPSO and the GA-based algorithms were conducted, and the approximate Pareto optimal set was derived from these 10 repetitions. All the experiments were performed on a personal computer with a Pentium IV 1.8 GHz CPU, 1GB RAM, and 80GB hard disk. The programs were coded with C Language.

Table 1 presents the values of the performance metrics for the competing algorithms used on the various-sized item banks. We observe that EMOPSO obtains the smallest values for the three metrics among the competing methods on all of the 10 item banks, meaning that the convergence and diversity of the non-dominated solutions produced by the EMOPSO are better than those obtained by using the GA-based and Random methods. The mean values of the three metrics are shown at the bottom of the same table. The ratios of the mean GD values for the GA-based and the
Random methods divided by the mean GD value for the EMOPSO method are 8.77 and 22.4, respectively, indicating that the non-dominated solutions produced by the EMOPSO method achieve better convergence than those generated by the other two methods. Analogously, the mean SP values show that the spacing variance of the non-dominated solutions obtained by EMOPSO is about 11 percent of that of the solutions produced by the GA-based method, so the EMOPSO generates solutions that have better diversity than those derived by the GA-based method. It is noteworthy that the Random method obtains fewer than three non-dominated solutions for the item banks with 250 and 500 items because it does not employ any learning mechanism and so easily gets trapped in the infeasible solution region when the size of the item bank is relatively small. Therefore, we are not able to compute the corresponding SP values. However, if we discard these two cases, the mean SP value for the Random method is less than that for the GA-based method (but still greater than that produced by the EMOPSO method). This is because the randomization helps the Random method search in different directions and improves diversity, but the convergence of the obtained solutions is not competitive when compared with those derived by the other methods. Finally, the ER values indicate that none of the solutions obtained by the GA-based and the Random methods is identified as a member in the approximate Pareto optimal solutions. On the other hand, an average of about 29 percent of the solutions generated in each run of the EMOPSO method belong to the set of approximate Pareto optimal solutions.

Table 1. Comparative performance values for the EMOPSO, GA-based, and the Random methods with various-sized item banks

<table>
<thead>
<tr>
<th>Item bank size</th>
<th>GD</th>
<th>SP</th>
<th>ER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EMOPSO</td>
<td>GA-based</td>
<td>Random</td>
</tr>
<tr>
<td>250</td>
<td>0.99E-3</td>
<td>2.56E-2</td>
<td>1.33E-2</td>
</tr>
<tr>
<td>500</td>
<td>1.28E-3</td>
<td>1.17E-2</td>
<td>6.58E-2</td>
</tr>
<tr>
<td>1000</td>
<td>0.77E-3</td>
<td>1.94E-2</td>
<td>2.80E-2</td>
</tr>
<tr>
<td>2000</td>
<td>0.92E-3</td>
<td>0.75E-2</td>
<td>1.96E-2</td>
</tr>
<tr>
<td>4000</td>
<td>1.63E-3</td>
<td>0.93E-2</td>
<td>1.73E-2</td>
</tr>
<tr>
<td>5000</td>
<td>1.32E-3</td>
<td>1.59E-2</td>
<td>2.12E-2</td>
</tr>
<tr>
<td>10000</td>
<td>1.02E-3</td>
<td>0.93E-2</td>
<td>2.49E-2</td>
</tr>
<tr>
<td>20000</td>
<td>0.92E-3</td>
<td>2.17E-2</td>
<td>1.23E-2</td>
</tr>
<tr>
<td>40000</td>
<td>7.02E-3</td>
<td>1.72E-2</td>
<td>1.50E-1</td>
</tr>
<tr>
<td>50000</td>
<td>1.31E-3</td>
<td>1.32E-2</td>
<td>3.19E-2</td>
</tr>
<tr>
<td>Mean</td>
<td>1.72E-3</td>
<td>1.51E-2</td>
<td>3.85E-2</td>
</tr>
</tbody>
</table>

*Not available because fewer than three non-dominated solutions were found.

Figure 5. Approximate Pareto front and the solution fronts produced by the competing methods
Figure 5 shows an instance of the approximate Pareto front and the solution fronts produced by the EMOPSO method and GA-based method for the 50,000-item bank. Four out of the 11 solutions produced by EMOPSO coincide with the approximate Pareto optimal solutions. By contrast, none of the seven solutions found by the GA-based method is a member of the approximate Pareto optimal set. In addition, the solutions obtained by EMOPSO are more uniformly distributed than those produced by the GA-based method. As a supplement, we implemented another solution method based on NSGA II using the program code in C for Windows provided at http://www.iitk.ac.in/kangal/codes.shtml. We observe that the NSGA II-based approach seems to produce solutions with better spacing than those obtained by the EMOPSO method or GA-based method, however, the GD quality (the capability of convergence to the Pareto optimal front) of the result generated by the NSGA II-based approach is worse than that obtained by the EMOPSO method.

**Sensitivity analysis**

To evaluate the robustness of the proposed method, we conducted sensitivity analysis on the 50,000 item bank by varying the value of the following problem parameters. The number of parallel test sheets (\(K\)) ranged from 2 to 4, the number of test items (\(S\)) to compose a test sheet was set to 20, 30, and 40, and the value of target difficulty (\(F\)) was specified as 0.7, 0.75, and 0.8, respectively. These parameter value variations can be practically applied to many real-world applications. Table 2 tabulates the values of the three performance metrics for the EMOPSO method and the GA-based method; the default values are \(K = 3, S = 30, \text{ and } F = 0.7\). We see that EMOPSO outperforms the GA-based methods on all test cases based on the three performance metrics. It is noteworthy that all of the solutions produced by the GA-based approach are dominated by those generated by our EMOPSO method.

<table>
<thead>
<tr>
<th>Parameter value variation</th>
<th>GD</th>
<th>GA</th>
<th>GD</th>
<th>GA</th>
<th>GD</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.32E-3</td>
<td>3.53E-2</td>
<td>3.41E-3</td>
<td>2.73E-2</td>
<td>0.69</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.31E-3</td>
<td>1.32E-2</td>
<td>2.60E-3</td>
<td>1.50E-2</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>7.40E-4</td>
<td>1.56E-2</td>
<td>1.66E-3</td>
<td>8.84E-3</td>
<td>0.59</td>
<td>1</td>
</tr>
<tr>
<td>(S)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.70E-3</td>
<td>4.21E-2</td>
<td>4.70E-3</td>
<td>2.85E-2</td>
<td>0.45</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>1.31E-3</td>
<td>1.32E-2</td>
<td>2.60E-3</td>
<td>1.50E-2</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>1.09E-3</td>
<td>3.12E-2</td>
<td>2.07E-3</td>
<td>6.39E-3</td>
<td>0.67</td>
<td>1</td>
</tr>
<tr>
<td>(F)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>1.31E-3</td>
<td>1.32E-2</td>
<td>2.60E-3</td>
<td>1.50E-2</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>0.75</td>
<td>3.51E-3</td>
<td>3.98E-2</td>
<td>6.44E-3</td>
<td>1.05E-2</td>
<td>0.71</td>
<td>1</td>
</tr>
<tr>
<td>0.8</td>
<td>4.15E-3</td>
<td>2.75E-2</td>
<td>8.54E-3</td>
<td>3.39E-2</td>
<td>0.66</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Comparative performance values of the EMOPSO method and GA-based method with various values of problem parameters

![Figure 6. Progressive solution fronts produced by EMOPSO at various numbers of generations](image-url)
Convergence behavior

We proceed by analyzing the convergence behavior of the EMOPSO method and GA-based method. Because both of the algorithms employed the same population size, the comparative convergence can be observed by showing the solution fronts produced when a given algorithm experiences various specified numbers of generations. Figures 6 and 7 show the progressive solution fronts obtained by the EMOPSO method and GA-based method at 100, 2,000, and 5,000 generations, respectively. Apparently, EMOPSO produces values of the first objective that are significantly better than those obtained by the GA-based method for each of the computational efforts. The values of the second objective manifest similar phenomena, though they are not as prevalent as those of the first objective. Overall, the EMOPSO algorithm converges more effectively towards the approximate Pareto front than does the GA-based method.

Questionnaire and analysis

To have a better understanding of the reaction of educators or potential users to this innovative approach, a questionnaire was distributed to 18 educators, including 6 elementary school teachers, 6 high school teachers and 6 college teachers, who had read the description of the parallel test sheet generating problem and experienced the testing system. The survey used a five-point Likert scale, with the responses coded as 1 = “strongly disagree” through to 5 = “strongly agree.” Table 3 shows the statistical results of the questionnaire.

Table 3. Statistical results of the questionnaire

<table>
<thead>
<tr>
<th>No.</th>
<th>Items</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The criteria in the problem model are reasonable.</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.50</td>
</tr>
<tr>
<td>2.</td>
<td>The objective function in the problem model reflects the actual needs of practical applications.</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.56</td>
</tr>
<tr>
<td>3.</td>
<td>The testing system is useful for holding large-scale or nation-wide tests.</td>
<td>12</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.67</td>
</tr>
<tr>
<td>4.</td>
<td>The testing system is efficient.</td>
<td>8</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4.33</td>
</tr>
<tr>
<td>5.</td>
<td>The system can reduce the labor burden for holding tests.</td>
<td>10</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4.50</td>
</tr>
<tr>
<td>6.</td>
<td>I would recommend this system to my colleagues or educational institute.</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4.44</td>
</tr>
<tr>
<td>7.</td>
<td>The use of icons and texts in this system is coherent.</td>
<td>2</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>3.78</td>
</tr>
<tr>
<td>8.</td>
<td>I can easily find the functions in the system.</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>3.56</td>
</tr>
<tr>
<td>9.</td>
<td>The interface design, including text color, size, and icons, is easy to identify.</td>
<td>0</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>3.33</td>
</tr>
<tr>
<td>10.</td>
<td>The messages provided by the system are concise.</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4.00</td>
</tr>
</tbody>
</table>
According to these results, most teachers think that the problem model is reasonable and reflects the actual needs of practical applications, showing important support for the development of this optimization algorithm and testing system. In addition, most of the teachers indicated that the testing system is useful for holding large-scale or nationwide tests, which meets the original objective of this study.

The statistical results also show several benefits of the innovative approach, including its efficiency and effectiveness (reducing the labor burden); moreover, 89 percent of the teachers surveyed would recommend the system to their colleagues or educational institutes.

**Concluding remarks and future work**

In this paper, we propose a particle swarm optimization (PSO) method to cope with the multi-objective parallel test-sheet composition (MOPTSC) problem, which involves multiple objectives and a set of complex constraints. To the best of our knowledge, this work is the first attempt to search for the Pareto optimal set in the domain of test-sheet composition, contrary to the traditional weighting sum method. It is of practical importance that the test-sheet designer would rather see a set of parallel test-sheet compositions of high quality than just a single instance, even though the difficulty of the former problem is surely greater than the latter. With these considerations in mind, we propose an enhanced multi-objective PSO (EMOPSO) by elaborating specific features involving constraint handling, restarting, and diversification strategies.

To verify the effectiveness and robustness of the proposed method, 10 test-item banks, each containing from 250 to 50,000 test items, were created. By experimenting with the 10 test-item banks, we compared the EMOPSO algorithm to the GA-based and random search methods. The numerical results show that the quality of the solutions produced by EMOPSO is better than that of the solutions derived by the GA-based and random search methods in terms of the three multi-objective metrics (GD, SP, and ER) that are widely used in the literature. In addition, the EMOPSO approach is more robust than the GA-based method under varying problem parameter values, and converges faster toward the approximate Pareto front.

For further applications, collaborative plans with several government agencies will be undertaken. In our future work, the approach will be used in several formal and large-scale tests, including national technical skill certification tests and language ability certification tests.

**Acknowledgement**

This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 96-2628-S-024-001-MY3 and NSC 97-2631-S-024-002.

**References**


Factors Influencing Adult Learners’ Decision to Drop Out or Persist in Online Learning

Ji-Hye Park¹ and Hee Jun Choi²

¹Department of Education, Kookmin University, Seoul, Korea // jpark22@kookmin.ac.kr
²Department of Education, Hongik University, Seoul, Korea // heejun.choi@gmail.com

ABSTRACT

The number of adult learners who participate in online learning has rapidly grown in the last two decades due to online learning’s many advantages. In spite of the growth, the high dropout rate in online learning has been of concern to many higher education institutions and organizations. The purpose of this study was to determine whether persistent learners and dropouts are different in individual characteristics (i.e., age, gender, and educational level), external factors (i.e., family and organizational supports), and internal factors (i.e., satisfaction and relevance as sub-dimensions of motivation). Quantitative data were collected from 147 learners who had dropped out of or finished one of the online courses offered from a large Midwestern university. Dropouts and persistent learners showed statistical differences in perceptions of family and organizational support, and satisfaction and relevance. It was also shown that the theoretical framework, which includes family support, organizational support, satisfaction, and relevance in addition to individual characteristics, is able to predict learners’ decision to drop out or persist. Organizational support and relevance were shown to be particularly predictive. The results imply that lower dropout rates can be achieved if online program developers or instructors find ways to enhance the relevance of the course. It also implies that adult learners need to be supported by their organizations in order for them to finish online courses that they register for.

Keywords

Adult dropout/retention/persistence, Online/distance learning, Theoretical framework for online dropout

Introduction

Distance learning allows adult learners who have employment, family, and/or other responsibilities to update knowledge and skills related to their job by saving travel costs and allowing a flexible schedule. Moore and Kearsley (2005) indicated that most distance education students are adults between the ages of 25 and 50. The number of programs for adult learners delivered online in corporate settings as well as in higher education has steadily increased over the last few years. According to the results of a survey administered by the National Center for Educational Statistics (NCES), 56 percent of all degree-granting higher education institutions offered distance courses during the 2000–2001 academic year (Waits & Lewis, 2003). In 2003, 34 percent of 1000 representative higher education institutions offered a complete online degree program (Allen & Seaman, 2004). In addition, Bersin (2005) indicated that online learning continued to grow in 2005 by 25 percent, and comprised 33 percent of all workplace learning. Sugrue and Rivera (2005) reported that training delivery via technology (or online) increased from 35 percent to 38 percent in large organizations, and from 24 percent to 27 percent in relatively smaller organizations from 2002 to 2003.

In spite of the growth in online learning, high dropout rates have been of concern to many organizations and higher education institutions. According to Meister (2002), 70 percent of adult learners enrolled in a corporate online program did not complete it. The Corporate University Xchange (2000) indicated that one of the difficult challenges of online programs is to retain learners. A number of studies have shown that a higher percentage of students participating in an online course tend to drop out compared to students in a face-to-face classroom (Hiltz, 1997; Phipps & Merisotis, 1999). Some consider the higher dropout rate in distance learning a failure while others advise careful interpretation of the issue because of unique characteristics and situations that online learners have. Diaz (2002) indicated that uncontrollable factors influence dropout decisions and a high dropout rate is not necessarily indicative of academic non-success. Nonetheless, it is still not easy to explain to corporate executives that dropout rates do not matter (Alexander, 2002), and it is certain that the issue of high dropout rates in online training should be addressed and dealt with.

Several theories and theoretical frameworks have been proposed to explain why students drop out. In particular, Tinto’s student integration model (1993) and Bean and Metzner’s student attrition model (1985) have guided dropout
research studies. Tinto (1993) claimed that attrition is a result of interactions between a student and his/her educational environment during the student’s stay in a program. He indicated that social integration and academic integration produced stronger student commitment to their institutions and increased students’ persistence. However, educators who desire to study the persistence of nontraditional students, who have different characteristics and nature from traditional students, have found that Tinto’s model has limited applicability (Rovai, 2003; Bean & Metzner, 1985). Tinto himself indicated that it was necessary to modify his model when used with nontraditional students (Tinto, 1982).

Bean and Metzner (1985) developed a conceptual model for nontraditional students who drop out that includes academic performance, intent to leave primarily influenced by academic and psychological outcomes, background and defining variables, and environmental variables. They asserted that the main difference between the attrition process of traditional and nontraditional students is that nontraditional students are more affected by the external environment than traditional students. However, Bean and Metzner’s model is unlikely to be applied for distance learners because there is a significant discrepancy between the definitions of distance learners in general and nontraditional students in the Bean and Metzner’s model (Kember, 1989).

Kember (1989), therefore, proposed a longitudinal process model of dropout distance education and made suggestions for testing the model (e.g., developing reliable instruments, conducting both qualitative and quantitative research, etc). Kember’s longitudinal model recognizes that social and academic integration of students should be viewed with intervening variables between initial student characteristics/background and persistence, that components change over time, and that students have to make dropout decisions several times during lengthy courses. Kember, Lai, Murphy, Siaw, and Yuen (1992, 1994) have tested this model in different sets of institutions, courses, and students and emphasized the importance of social and academic integration to student progress in distance learning. Since then, a couple of researchers have committed to comparing those previous models, determining advantages and disadvantages, and finally developing a model explaining the process of dropping out in a particular population and learning environment. Cabrera, Castaneda, Nora, and Hengstler (1992) reviewed Tinto’s and Bean and Metzner’s dropout frameworks, and the results indicated that Tinto’s model is more comprehensive and robust while Bean and Metzner’s model accounts for more variance in persistence.

Rovai (2003) proposed a persistence model to explain factors affecting a learner’s decision to drop out of online learning. The model included two prior-to-admission variables and two after-admission variables. The two prior-to-admission variables are student characteristics and student skills prior to admission. The two after-admission variables are external factors (e.g., finances, hours of employment, outside encouragement, etc.) and internal factors (e.g., academic integration, social integration, self-esteem, interpersonal relationships, study habits, advising, absenteeism, etc.). Rovai’s framework is established by a thorough review of the most comprehensive previous frameworks (i.e., Tinto’s student integration model [1993] and Bean & Metzner’s student attrition model [1985]), particularly focusing on nontraditional online learners who have characteristics similar to adult learners in organizations. This model was also tested and expanded by Packham, Jones, Miller, and Thomas (2004).

Park (2007) reviewed studies that focused on identifying factors affecting non-traditional and non-degree online program students who drop out and proposed a framework based on Rovai’s model for understanding adult dropouts (Figure 1). Based on the review, she indicated that the significance of the four factors from Rovai’s model is supported from many studies with a variety of research methods. However, she suggested revision of the structure of the model and elimination of some of the variables. Specifically, learner skills are in a grey box because these have found little empirical support in previous studies, and their inclusion can be determined only through relevant further investigation. The external factors are moved between prior to and during the courses because these affect student decisions not only during the course but also prior to the course. Adult distance learners may drop out of the course due to increased workload or job change that happens during the course, but some learners may drop out of a course even before they start because of such external reasons. In addition, external factors and internal factors are likely to interact with each other. For example, when learners have a heavy workload and little time for study, they are more likely to drop out of a course when they cannot get feedback or if it is hard to contact the instructors than when they can easily communicate with them and get more responses. If proper course design and technology are being used, some external problems are likely to be mitigated. So the relationship between internal factors and external factors are expressed as inter-correlation rather than as a one-sided influence. In addition, it appeared that only internal factors would have a direct influence on persistence decision, and others have an indirect affect through internal factors in Rovai’s model. However, many studies have reported that some external factors have been major reasons
why online learners decided to drop out, particularly in relation to adult distance learners. Therefore, a direct line from external factors to dropout/persistence has been added.

Figure 1. Theoretical framework for adult dropout in online learning (Park, 2007)

Even though numerous studies have tried to identify factors affecting learners’ decision to drop out, only a dozen research studies have empirically explored this issue, and no consensus has been reached for which factors have definite influences on the decision (Park, 2007). Although Park and Choi (2007) presented a research study investigating the effects of individual characteristics, external factors, and internal factors on non-traditional adult learners’ decision to drop out, the study was limited in that the sample size was too small ($n = 47$) so that the results could hardly be reliable. In addition, relevance, one of the most crucial motivational factors considered to affect adult learners’ decision to drop out, was omitted.

Accordingly, the purpose of this study was to identify meaningful factors affecting learners’ decision to drop out of online courses and ultimately to shed light on how we can retain students in online courses by involving a significant number of research participants and adding meaningful factors. Particularly, this study focused on the three main categories: individual characteristics, external factors, and internal factors. To be more specific, age, gender, educational background, and employment status were chosen as individual characteristics because these four are the most often cited factors in previous studies (Park, 2007). External factors consist of family support and organizational support. Most adult learners have many responsibilities for their family as well as for their job, and these two are key factors affecting adult learners’ decision to drop out of online courses (Park, 2007). Motivation is one of the most frequently studied variables in relation to dropout (Chyung, 2001; Chyung, Winiecki, & Fenner, 1998; Doo & Kim, 2000; Jun, 2005; Levy, 2007; Menager-Beeley, 2004). In particular, relevance and satisfaction are the sub-dimensions of motivation that have frequently been studied (Chyung et al., 1998; Doo & Kim, 2000; Levy, 2003, 2007; Shea, Pickett, & Pelz, 2003) and are known to be highly correlated with various course-related issues such as instructional design, organization of the online courses, instructors’ facilitation, and interaction (Shea et al., 2003). This study could not include other internal factors such as social integration, academic integration, and technology issues shown in the above framework because the courses investigated in this study were developed before conducting this study, and the researchers did not have access to the course contents and were not involved with the design and development process.

Therefore, the research questions used in this study were as follows:

1. Do the dropouts and the persistent learners of online courses show differences in their individual characteristics (i.e., age, gender, and educational background), external factors (i.e., family support and organizational support), and internal factors (i.e., motivation in terms of satisfaction and relevance)?
2. What factors are significant to predict learners’ decision to drop out of online courses?

**Methods**

**Population and sample**

The target population of this study was non-traditional adult learners who enroll in job-related online courses offered by a large Midwestern university. Since November 2002, 18 distance courses were offered, with 378 learners registering and 204 learners completing (dropout rate = 46.0%). Since 2005, when the current learning system was established, three online courses were offered three times, and 107 out of 234 participants completed the courses (dropout rate = 54.2%). The increasing dropout rate has become a concern to the program providers as well as to many other distance programs and institutions.

The sample for this study was 147 adult learners who have either completed or dropped out of one of the three online courses offered between fall 2005 and summer 2007 and who agreed to participate in this study. Of the 147 participants, 98 (66.7%) were persistent learners while 49 (33.3%) were dropouts; 105 (71.4%) were female while 42 (28.6%) were male. About half of the participants were between 30 and 39 years old (73 participants, 49.7%) and had some college education (79 participants, 53.7%), as shown in Table 1.

<table>
<thead>
<tr>
<th>Demographic variables</th>
<th>Persistent</th>
<th>Dropout</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>73 (74.5)</td>
<td>32 (65.3)</td>
<td>105 (71.4)</td>
</tr>
<tr>
<td>Male</td>
<td>25 (25.5)</td>
<td>17 (34.7)</td>
<td>42 (28.6)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–29</td>
<td>14 (14.3)</td>
<td>13 (26.5)</td>
<td>27 (18.4)</td>
</tr>
<tr>
<td>30–39</td>
<td>53 (54.1)</td>
<td>20 (40.8)</td>
<td>73 (49.7)</td>
</tr>
<tr>
<td>Over 40</td>
<td>31 (31.6)</td>
<td>16 (32.7)</td>
<td>47 (32.0)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>37 (37.8)</td>
<td>17 (34.7)</td>
<td>54 (36.7)</td>
</tr>
<tr>
<td>College</td>
<td>50 (51.0)</td>
<td>29 (59.2)</td>
<td>79 (53.7)</td>
</tr>
<tr>
<td>Graduate or higher</td>
<td>11 (11.2)</td>
<td>3 (6.1)</td>
<td>14 (9.5)</td>
</tr>
<tr>
<td>Class</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9 (9.2)</td>
<td>12 (24.5)</td>
<td>21 (14.3)</td>
</tr>
<tr>
<td>2</td>
<td>38 (38.8)</td>
<td>16 (32.7)</td>
<td>54 (36.7)</td>
</tr>
<tr>
<td>3</td>
<td>51 (52.0)</td>
<td>21 (42.9)</td>
<td>72 (49.0)</td>
</tr>
</tbody>
</table>

NOTE (%)

**Instrumentation**

Family support and organizational support were measured by six-item, five-point Likert scales. Sample items from the family support measure are “My family understands me very well even though I spend little time with them because of online courses” and “My family is proud of me when I learn to improve my job performance.” Sample items from organizational support measure are “My organization is willing to reduce my workload when I need training for my job” and “My supervisor shows interest in my job-related learning.” The measures were developed by the researchers and reviewed by another researcher who has expertise in measurement. These were also pilot tested with former learners who finished online courses offered from the same department. Consequently, Cronbach’s alpha scores for family and organizational support were .78 and .83, respectively.

Items used to measure satisfaction and relevance were originally based on Keller’s (1987) Instructional Materials Motivation Survey (IMMS). This study did not use Keller’s Course Interest Survey (CIS), which is also designed to measure motivation in a specific course because the IMMS has been extensively examined for reliability while the CIS has been used by only a few studies (Gabrielle, 2003). Word choice and the number of items were slightly different from those in the original IMMS because of the need to be more specific for this study. Sample items to measure satisfaction are “It was a pleasure to work on this type of instruction” and “I want to receive this type of instruction again for my job improvement.” Previous studies found that these items had a Cronbach’s alpha of 0.86 (Choi & Johnson, 2005 & 2007). This study found Cronbach’s alpha of these items is .90. Sample items to measure
relevance are “The content of this course is relevant to my interest,” “There are explanations and examples of how people use the knowledge,” and “I could relate the content of this course to things that I have seen, done, or thought about in my own life.” Cronbach’s alpha of these items obtained from this study is .87.

Data collection and analysis

Quantitative data from a survey questionnaire consisted of learners’ age, gender, and educational level, their perceptions of family support and organizational support, and motivation in terms of satisfaction and relevance. Participation in this study was solicited and data were collected via email because of participants’ geographic locations. The researcher sent an email to the 234 participants registered to introduce the purpose of this study and to solicit their participation. After sending the message three times, 161 learners showed interest in participating in this study. The researchers then sent an online survey to them. Finally 149 submitted responses, but only the data from 147 could be used for this study because two persons did not answer most of the questions.

The answers to the research questions for this study were reached using several statistical tools. The data to answer the first research question were analyzed by descriptive statistics, chi-square, and multivariate analysis of variance (MANOVA). Chi-square was used to determine the differences in the individual characteristics between the two groups in this study; that is, dropouts and persistent learners. MANOVA is used to analyze a dependence relationship represented as the differences in a set of dependent measures across a series of groups (Hair, Black, Babin, Anderson, & Tatham, 2006). In this study it was used to determine the differences between the two groups because this study used multi-dependent variables. Two-way factorial design was employed to control the effect of the course as well when examining group differences in family support, organizational support, and motivation in terms of satisfaction and relevance. Learners may show different perceptions and/or motivation depending on the course for which they register. The second research question was answered by logistic regression analysis. This is a specialized form of regression and is formulated to predict and explain a two-group categorical variable (Hair et al., 2006).

Results

Research question 1

Table 1 shows the individual characteristics of persistent and dropout learners. Females made up 74.5 percent of persistent learners and 65.3 percent of the dropouts. A significant portion of learners ranged from 30 to 39 years of age (54.1% and 40.8%, respectively) and had a college degree (51.0% and 59.2%, respectively) in both groups. Even though more learners in the dropout group were between 20 and 29 years old (26.5%) than in the persistent group (14.3%), the difference was not statistically significant ($\chi^2 [2, 147] = 3.84, p = .147$). Overall, statistical significances in differences of individual characteristics were not found ($\chi^2 = 1.35 \sim 3.84, p = .147 \sim .501$).

Table 2 presents perceptions of family support, organizational support, satisfaction, and relevance by group. Overall, the persistent group showed higher means in perceptions of family and organizational support and higher means in satisfaction and relevance than the dropout group. Table 2 also showed the means of perceptions of dependent variables by class. Learners who registered for the first class showed the biggest mean differences between the two groups. In the first class, the mean differences in the four dependent variables between the two groups were 5.93, 9.30, 7.11, and 5.78 respectively. The mean differences in the second and third classes were smaller than those in the first class.

<table>
<thead>
<tr>
<th>Dependant variables</th>
<th>Class</th>
<th>Persistent group</th>
<th>Dropout group</th>
<th>Mean differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Family support</td>
<td>1</td>
<td>9</td>
<td>27.22</td>
<td>5.47</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>38</td>
<td>26.62</td>
<td>4.99</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>51</td>
<td>26.49</td>
<td>5.11</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>98</td>
<td>26.61</td>
<td>5.04</td>
</tr>
</tbody>
</table>
The results from MANOVA are presented in table 3. The interaction effects of satisfaction and relevance across groups by class were statistically significant, and $F(2, 147) = 3.58, p < .05$, and $F(2, 147) = 3.62, p < .05$, respectively. As shown in figures 2 and 3, the results in the first class showed larger mean differences between the two groups, and this observation could be confirmed from Table 2. The interaction effect is deemed ordinal because the lines depart from parallel but never cross in a significant amount (Hair et al., 2006). Because the interaction effect is deemed significant but ordinal, the analysis of the main effects for the differences across the groups could proceed. The main effects of the group on the dependent variables were statistically significant ($F[1, 147] = 11.82, p < .001$, $F[1, 147] = 87.70, p < .001$, $F[1, 147] = 54.77, p < .001$, and $F[1, 147] = 58.70, p < .001$, respectively). In other words, dropouts and persistent learners perceived differently the level of family support and organizational support that they had received and showed different levels of motivation in terms of satisfaction and relevance. Lastly, the main effects of the class on the dependent variables were analyzed. The results indicated that learners showed different levels of satisfaction and relevance according to the class that they registered for ($F[2, 147] = 35.24, p < .001$, and $F[2, 147] = 36.49, p < .001$, respectively). In sum, the results showed that the persistent group and the dropout group showed differences in their perception of family support and organizational support and in the level of motivation in terms of satisfaction and relevance. In addition, learners showed different levels of motivation depending on the class for which they registered.

<table>
<thead>
<tr>
<th>Group (dropout/persistent)</th>
<th>Family</th>
<th>Organization</th>
<th>Satisfaction</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type III SS</td>
<td>339.87</td>
<td>1711.84</td>
<td>514.99</td>
<td>351.87</td>
</tr>
<tr>
<td>df</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MS</td>
<td>339.87</td>
<td>1711.84</td>
<td>514.99</td>
<td>351.87</td>
</tr>
<tr>
<td>$F$</td>
<td>.000***</td>
<td>.001***</td>
<td>.000***</td>
<td>.000***</td>
</tr>
<tr>
<td>$p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Family</th>
<th>Organization</th>
<th>Satisfaction</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type III SS</td>
<td>23.52</td>
<td>22.76</td>
<td>662.73</td>
<td>437.57</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>MS</td>
<td>11.76</td>
<td>11.38</td>
<td>331.36</td>
<td>218.79</td>
</tr>
<tr>
<td>$F$</td>
<td>.41</td>
<td>.58</td>
<td>.3524</td>
<td>.3649</td>
</tr>
<tr>
<td>$p$</td>
<td>.560</td>
<td>.665</td>
<td>.000***</td>
<td>.000***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group * Class</th>
<th>Family</th>
<th>Organization</th>
<th>Satisfaction</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type III SS</td>
<td>59.34</td>
<td>13.98</td>
<td>67.41</td>
<td>43.45</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>MS</td>
<td>29.67</td>
<td>6.99</td>
<td>33.70</td>
<td>21.73</td>
</tr>
<tr>
<td>$F$</td>
<td>1.03</td>
<td>.36</td>
<td>3.58</td>
<td>3.62</td>
</tr>
<tr>
<td>$p$</td>
<td>.700</td>
<td>.359</td>
<td>.030*</td>
<td>.029*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error</th>
<th>Family</th>
<th>Organization</th>
<th>Satisfaction</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type III SS</td>
<td>4055.49</td>
<td>2752.28</td>
<td>1325.94</td>
<td>845.21</td>
</tr>
<tr>
<td>df</td>
<td>141</td>
<td>141</td>
<td>141</td>
<td>141</td>
</tr>
<tr>
<td>MS</td>
<td>28.76</td>
<td>19.52</td>
<td>9.40</td>
<td>5.99</td>
</tr>
</tbody>
</table>

*$p < .05$, **$p < .01$ ***$p < .001$
Research question 2

Logistic regression analysis was used to determine how well the four variables (family support, organizational support, satisfaction, and relevance) could predict learners’ decisions to drop out, identify individual characteristics, and determine which factors are most significant to explain the decision. Individual characteristics were entered in the first step because they are what learners bring with them to the class. Other variables measuring learners’ perceptions of family support and organizational support and motivation in terms of satisfaction and relevance then were entered in the second step in order to determine how these variables predict learners’ decision to drop out or persist beyond their individual characteristics.
Table 4 presents the predictability of the proposed model. The result showed that the model correctly classified 90.8 percent of the persistent learners and 87.8 percent of the dropout learners, for an overall accuracy of 89.8 percent and is acceptable to predict adult learners’ decision to drop out of online courses ($\chi^2[11, 147] = 111.59, \ p < .001$, Nagelkerke $R^2 = .74$).

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Persistent</th>
<th>Dropout</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>89</td>
<td>9</td>
<td>90.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>43</td>
<td>87.8</td>
</tr>
<tr>
<td>Overall percentage</td>
<td></td>
<td></td>
<td></td>
<td>89.8</td>
</tr>
</tbody>
</table>

Table 5 represents the logistic regression coefficient (B), standard error (S.E.), Wald test, and odds ratio (Exp [B]) for each of the predictors. When entering demographic information, including the class for which learners register in the first step, there was no significant predictor. In the second step, however, organizational support and relevance turned out to be significant predictors in addition to the class. The class was insignificant in the first step and turned out to be significant in the second step because of the interaction between the class and the group. Based on the results obtained here, it can be concluded that learners’ perceptions of organizational support and relevance are significant predictors to learners’ decision to drop out or persist in an online course. It can be interpreted that online learners’ decision to drop out or persist is related to the class for which they register. It also implies that learners are more likely to drop out of an online course when they perceive that the organization does not support their learning and the course is not related to their own lives.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>S.E.</td>
</tr>
<tr>
<td>Age</td>
<td>1.405 (2)</td>
<td>.495</td>
</tr>
<tr>
<td>Age (1)</td>
<td>.219</td>
<td>.546</td>
</tr>
<tr>
<td>Age (2)</td>
<td>-.369</td>
<td>.457</td>
</tr>
<tr>
<td>Gender</td>
<td>-.347</td>
<td>.401</td>
</tr>
<tr>
<td>Education</td>
<td>1.628 (2)</td>
<td>.443</td>
</tr>
<tr>
<td>Education (1)</td>
<td>-.456</td>
<td>.742</td>
</tr>
<tr>
<td>Education (2)</td>
<td>.333</td>
<td>.392</td>
</tr>
<tr>
<td>Class</td>
<td>3.013 (2)</td>
<td>.222</td>
</tr>
<tr>
<td>Class (1)</td>
<td>.872</td>
<td>.573</td>
</tr>
<tr>
<td>Class (2)</td>
<td>-.098</td>
<td>.440</td>
</tr>
<tr>
<td>Family support</td>
<td>.054</td>
<td>.071</td>
</tr>
<tr>
<td>Organizational support</td>
<td>-.461</td>
<td>.103</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>-.049</td>
<td>.116</td>
</tr>
<tr>
<td>Relevance</td>
<td>-.566</td>
<td>.575</td>
</tr>
</tbody>
</table>

*p < .05, ** p < .01 ***p < .001

Conclusions

This study aimed to determine whether or not there were differences between dropouts and persistent learners in online courses in their individual characteristics, the perceptions of family support and organizational support, and the level of motivation (i.e., satisfaction and relevance). This study also intended to find factors to predict dropouts and persistent learners in online courses to help stakeholders associated with online courses for adult learners find ways to lower the high dropout rates.
Previous studies have reached no consensus on the influence of learner characteristics on adult learners’ decision to drop out of online courses. Some reported that learner characteristics have significant influence on the decision (Brown, 1996; Jun, 2005; Meneger-Beeley, 2004; Osborn, 2001; Packham et al., 2004), while others claimed those characteristics have only minor or indirect effect (Kember et al., 1992, & 1994; Willging & Johnson, 2004). This study added additional evidence for the latter by showing that the persistent learners did not differ from the dropouts in their individual characteristics. In other words, learners’ age, gender, and educational level did not have a significant and direct effect on the dropout decision. Although the result does not claim that individual characteristics should be ignored, it can be concluded that individual characteristics have little influence on the decision to drop out and thus can be considered as trivial.

External factors such as organizational supports, financial problems, and time constraints have been known to be crucial obstacles to adults’ participation in learning because adult learners are associated with various roles in their lives (Darkenwald & Merriam, 1982; Johnstone & Rivera, 1965). It is true not only in traditional adult learning programs but also in online learning programs. Willging and Johnson (2004) claimed that external factors such as family issues, lack of organizational support, changing job, and workload are the main factors affecting the decision to drop out of online courses. Greer, Hudson, and Paugh (1998) emphasized family and peer support for success of online learners. Rovai (2003) also emphasized the effect of nonschool factors that conflict with academic life on students’ decision to drop out. Although these studies have already claimed the influence of external factors on the dropout decision, these studies are limited in that the first two used only a limited number of subjects and the third one was based only on conceptual analysis. Accordingly, the results were hardly generalizable to learners in different environments, and additional empirical evidence was needed to support the contention. This study showed that dropouts were significantly different from persistent learners in external factors (i.e., family support and organizational support), and the results of this study are consistent with those of previous studies.

Adult learners are more likely to drop out of online courses when they do not receive support from their family and/or organization while taking online courses, regardless of learners’ academic preparation and aspiration. Internal factors such as course design strategies and learners’ motivation should be prioritized at the course development stage in order to make the course participatory and interesting and to keep learners engaged. Once the course is launched and in progress, however, course administrators and instructors should consider external factors that might interrupt learners’ participation and persistence. It would be difficult for them to control the external factors. However, it is important to consider learners’ situation while managing or maintaining the course so that learners can get help if needed. In the event that an instructor knows that learners are not receiving enough support from their family and organization, he/she might help the learners stay in the course by paying extra attention, using appropriate motivational strategies, and providing additional internal support. In addition, course administrators and instructors need to inform learners’ family and organization of the advantages of the course in order to induce their supports. This study concludes that instructional designers should systematically analyze external factors surrounding learners and use the analysis results to initiate learning and motivate learners so that the high dropout rate can be decreased. The results of this study support the significance of those external factors that are easily overlooked by instructional designers.

This study also showed that dropouts had significant differences in perceptions of learner satisfaction and relevance from persistent learners. In other words, learners are less likely to drop out when they are satisfied with the courses, and when the courses are relevant to their own lives. In agreement with prior research (e.g., Levy, 2007; Doo & Kim, 2000), the results suggest that learners’ satisfaction with the online course and relevance to learners’ job, prior knowledge, and experiences are major factors affecting their decision to drop out or persist. From this result, a course designer or instructor might get insight about how to design the course better. In order to enhance satisfaction as a way to motivate online learners, rewards such as a completion certificate, praise, and promotion should be given to learners. By providing opportunities to apply newly acquired knowledge into real situations, learners can feel that the skills and knowledge obtained from the course are useful and satisfactory and thus they can be motivated to persist in the course. Relevance can be achieved by designing a course that contains learning materials and cases closely related to learners’ interests, experiences, goals, and so forth. Keller (1987) suggested that relevance could be established by using learners’ experiences, allowing learners to choose learning methods and strategies, and meeting learners’ expectations and goals. Online learners can easily lose motivation unless the course is designed to stimulate their active participation and interaction and meet their expectations. Therefore, an online course needs to be designed in ways to guarantee learners’ satisfaction and be relevant to learners’ needs.
The result from the logistic regression analysis reported the particular importance of organizational support and relevance, in addition to the class for which they registered. Organizational support and relevance were statistically significant predictors of learners’ decision to drop out or persist in online courses. The result implies that learners are more likely to decide to drop out when they are not getting organizational support for their learning. Often adult learners should be granted released time from their jobs and given encouragement from peers and/or supervisors to participate in the course. It is apparent that organizational support could be a crucial factor to influence learners’ dropout decision. The result also implies that learners who perceive that the course is relevant to their job or life are less likely to make a decision to drop out. Adult learners tend to prefer learning that has a practical purpose to learning for academic purposes only. Therefore, online courses should be designed in ways to allow adult learners to apply their learning to their real lives.

The sample of this study was selected from only one institution in the U.S. Thus, the results from this study may not be generalizable to adult learners in other institutions and/or countries. Further investigation is needed to confirm the generalizability of the results to broader populations. This study includes a limited number of variables even though they were chosen for their importance based on thorough review of the literature. There are many other variables, including two other motivation factors (i.e., confidence and attention), and factors associated with instructional strategies that may affect adult learners’ decision to drop out of online learning. Further research, therefore, is needed to involve additional relevant factors and to expand the model to better explain and predict adult learners’ decision to drop out of online courses.

Acknowledgement

This study was supported by the 2008 Hongik University Research Fund.

References


Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. Internet and Higher Education, 6, 1–16.


A Five Dimensional Model for Educating the Net Generation

Ronald Noel Beyers
Young Engineers and Scientists of Africa Programme, Meraka Institute, Pretoria, South Africa // ron@yesa.org.za

ABSTRACT
This paper proposes a multi-dimensional concept model of an ICT enabled classroom to highlight potential similarities and differences between where teachers perceive themselves relative to their learners. Some teachers and learners may be at the two dimensional text-book level, while others are operating in at a globalization level. Being armed with this type of information, educators may be empowered to rethink their educational strategies especially where ICTs are employed in an educational context.

Keywords
Net generation, Model of education, Knowledge society, Digital divide, Globalization

Introduction

The world in which our children live is significantly different from that of yesterday. Today’s youngsters use laptops, pagers, instant messaging, and cell phones to connect to friends, family, experts, and others in their community and around the globe. They are bombarded with visual messages from the media – messages specifically targeted to tap into the billions in discretionary spending they control and/or influence. This is a generation that expects to actively participate in and through their media, hence the decrease in time spent by teens in viewing television and the corresponding increase in time spent on computers, gaming, and the Internet. Our children now have at their fingertips a virtual world – with all its promises and pitfalls (Lemke, 2003:5).

Schooling today is an attempt to make mini-scholars out of students by giving them doses of what was meant by scholarship in the eighteenth and nineteenth centuries. The theory of knowledge implicit here is that the educated person knows something about all the great books. This idea works when there aren't that many books in print (or at least it seems to work). But in an age when no one could possibly know something from every book that has been written, when there are enough books to go around, and when there are so many other forms of knowledge available to students, these ideas are outmoded. We must look to concepts that relate to today's world, one where there's so much to know that it is likely that students will have to direct their own education out of practical necessity (Schank & Cleary, 2008).

Public education is struggling to adapt to an intellectual, social and cultural transformation that has begun to emerge during the last thirty or forty years. New understandings on the frontiers of science, a growing awareness of the threats to planetary ecology, and a disruption of local communities and economies by the rise of globalization have made it necessary to rethink many of the basic assumptions that guided the development of modern industrial culture. It is increasingly evident that humanity faces the task of moving from an age of modernity into an uncharted post-industrial or post-modern future (Miller, 2000).

Figure 1. presents a diagrammatic representation of the relationship between an education system and on society in general. The formative years of schooling are aimed at preparing learners for an exit point into society after they successfully complete up to twelve consecutive grades. A mechanistic view of the education system demands that a feedback loop is essential to ensure that the needs of society are met through the system and is embedded in the contents of the formal curriculum which is managed by the government of the day. Change in society is inevitable resulting in a demand for learners with a new set of skills implying the need for these changes to be reflected in the curriculum of the day.

The curriculum in any contemporary democratic society always reflects the definition of democracy which that society has accepted as legitimate and true. Similarly, attempts to challenge the validity and legitimacy of a society’s dominant definition of democracy always find expression in attempts to challenge the form and content of the curriculum of that society’s schools. In this sense, the debates about the curriculum that occur in a democracy at any given time will reveal both how that democracy interprets itself, and how that interpretation is being challenged and revised in order to bring into being a more genuinely democratic form of life than that which currently exists. The curriculum in a democracy is thus always a curriculum for democracy, incorporating both a record of its past and a message for its future (Carr, 1998).
One of the primary functions of an education system is to convey and ensure a mastery of a set of knowledge, skills, attitudes and values that a particular society regards as desirable. During the formative years an individual will also be exposed to a cultural framework, which will supplement their survival strategies for the rest of their lives in a given society. One of the problems that many societies are facing especially where there is a transition from an industrial and manufacturing based economy to a knowledge society, is the rate of change. The reform processes in education, as portrayed through the curriculum, are seldom able to keep pace with the change resulting in students exiting the formal education system being inadequately prepared as the world that they were prepared for may no longer exist. To address this problem it is proposed that curriculum changes need to be visionary and project the activities within the classroom to intersect with the future needs of society as portrayed in Figure 1.

Teachers in the classrooms of today are facing an educational dilemma. The world that they grew up in and were trained in is also rapidly evolving around them. In a third world country like South Africa, these permutations are even more exaggerated by such a diverse variety of schools ranging from classrooms under trees with no running water, toilets or electricity in deep rural areas to modern digital classrooms with laptops and campus-wide radio networks found in the more affluent private institutions. The apartheid era has left a legacy which will take generations to rectify leaving many learners with very little chance of crossing the digital divide. The net effect for these individuals is a life of economic misery where they are excluded from many benefits that a technological world has to offer.

On the other hand ‘wiring the schools and populating them with computers is necessary but insufficient to ensure equal opportunity to share in the digital revolution’ (Tapscott, 1988:262). They need a redesigned education system and teachers who have been retrained and reoriented. Innovative technologies cannot make up for educational professionals who lack innovative methods and merely replicate learning models that don't work (Hooper, 2002).

A lack of sustained and meaningful professional development for teachers coupled to a stagnant curriculum can leave learners uninspired and unmotivated to extend themselves. This phenomenon is common as some teachers may not have, for whatever reason, had any in-service training or been exposed to new teaching methodologies and a phenomenon of ‘ignorance is bliss’ is perpetuated. The learners at the end of the day loose out on the benefits that educational reform has to offer such as the recently introduce Outcomes Based Education in the fledgling democracy of the South African education system.

A learner growing up in a first world environment has a very good chance of being nurtured on a digital diet acquiring the necessary skills to cope with life in a technological world. Children may be developing greater digital literacy than siblings who are just a few years older. For example, over two million American children (ages 6–17) have their own Web site. Girls are more likely to have a Web site than boys (12.2 percent versus 8.6 percent) (Oblinger & Oblinger, 2005). Individuals raised with the computer deal with information differently compared to previous cohorts: they develop hypertext minds, they leap around (Prensky, 2001).
Learning Theories

Gone are the days of working with the unformed, featureless minds of learners of the philosopher of John Locke’s ‘tabula rasa’. Pavlov’s classical conditioning as part of behaviourism together with Thorndike, Watson and particularly Skinner’s operant conditioning have been superseded by Jean Piaget’s constructivism. Though their relevance has been surpassed by constructivism previous philosophical underpinnings laid the foundation for the new paradigm to understand the nature of learning.

Constructivism is a philosophy of learning which articulates mechanisms by which knowledge is internalized by learners. Constructivism is an epistemological view of knowledge acquisition emphasizing knowledge construction rather than knowledge transmission and the recording of information conveyed by others. The role of the learner is conceived as one of building and transforming knowledge (James, Applefield, & Mahnaz, 2001).

Constructivism has multiple meanings. In philosophy it is an epistemological theory, in cognitive psychology it describes human learning, and in curriculum it is a set of design principles. Constructivism is an overarching theory that can incorporate a number of teaching practices such as co-operative, collaborative and inquiry-based learning…

Constructivism is basically a theory based on observation and scientific study about how people learn. It says that people construct their own understanding and knowledge of the world through experiencing things and reflecting on those experiences. When we encounter something new we have to reconcile it with our previous ideas and experience, may be discarding the new information as irrelevant. We are active creators of our own knowledge (Mechlova, Konicek, & Balnar, 2003).

Constructivists believe that knowledge is constructed and learning occurs when children create products or artefacts. They assert that learners are more likely to be engaged in learning when these artefacts are personally relevant and meaningful (Bhattacharya & Han, 2001). In a more traditional teacher-centric classroom where information transfer is prevalent learners are less likely to be engaged as the meaning is being interpreted for them by the teacher. Modern youth often find themselves immersed by a range of technological devices that provide them with access to information through a broad range of media. The challenge for the teacher is to capitalize on this by engaging the in the content of the curriculum through the use of the ‘tools’ of the Net Generation.

The Net Generation is defined as the population of about 90 million young people who have grown up or are growing up in constant contact with digital media (Tapscott, 1988). Born between roughly 1980 and 1994, the Millennials have already been pegged and defined by academics, trend spotters, and futurists: They are smart but impatient. They expect results immediately. They carry an arsenal of electronic devices - the more portable the better. Raised amid a barrage of information, they are able to juggle a conversation on Instant Messenger, a Web-surfing session, and an iTunes playlist while reading Twelfth Night for homework. Whether or not they are absorbing the fine points of the play is a matter of debate.

Millennials expect to be able to choose what kind of education they buy, and what, where, and how they learn. To meet the demands of these new students, they say, colleges must rethink how they operate. Imagine classrooms that incorporate more videos and video games, classes that meet electronically to fit students’ schedules, students who choose to learn from each other rather than a professor, and courseware, search engines, and library databases that are animated, image-based, and interactive (Carlson, 2005).

By contrast many practicing teachers grew up without computers and the chances of them being formally trained to use them in the classroom as slim. Teacher training programmes are having to be implemented for both in-service and teacher training institutes to bring them up to speed.

A Proposed Conceptual Model

It is easy to criticize an education system which is inherently about learning but at the same time it is critical to look at the teachers who form the engine room of the system. This proposed Five Dimensional Model suggests a contextualization of possible differences between the traditional teachers in the classrooms who may be unaware of the needs of the net generation that they encounter. The model proposes a multi-dimensional perspective of
comprehending possible similarities and differences which could be used to empower teachers to make greater use of technology in the classrooms while addressing the needs of the learners.

These learning theories fall short of encapsulating the complexities of the net generation. The need is to present a model which may provide for a differentiation of a multitude of dimensions which is indicative of the ability of the modern learner to operate on a number of different levels and their ability to multitask within such an environment. This model is a new approach and must not be seen as being in conflict with such theories as constructivism but rather supplements them. The model was based on a number of years of exposure to innovative learners including a variety of information communication technologies (ICTs) and educational pedagogies ranging from drill and practice to blended learning, teacher training and e-learning opportunities. It is envisaged that the model could provide educators with the means to comprehend the ‘differentness’ and ‘complexity’ of the net generation. A tool to graphically plot individuals within this 5D model is being developed which will serve to highlight similarities and differences between teachers and pupils especially where ICTs are deployed in the classroom. The intention is to sensitize teachers to this information so that they are encouraged to grapple with these issues to better understand them for the benefit of their learners.

The prevailing technologies of a particular place and time have always been intimately linked with education, because a society’s tools are both the subject and the means of its learning. Today, the fact that technology pervades almost every sphere of life – from home to work to play – results in profound implications for learning, both in schools and throughout life (Partnership for 21st Century Skills, 2007:6).

The challenge of evolving pedagogy to meet the needs of Net-savvy students is daunting, but educators are assisted by the fact that this generation values education. These students learn in a different way than their predecessors did, but they do want to learn (Barnes, Marateo, & Ferris, 2007).

This five-dimensional model of viewing an educational system suggests that education is no longer a linear process of information transfer, but rather that it has dimensions which can range from two dimensional textbook scenarios as well as a time and space dimension where ICTs are concerned. Each of the five dimensions have a predefined characteristics. The model is a mental picture that helps us understand something we cannot see or experience directly (Dorin, Demmin, & Gabel, 1990).

The First Dimension – X Axis (Survival Strategies)

This dimension is characterized by the hunter-gatherer leading to the period of farming which brought about an improvement in the quality of life during the agricultural era. The young were trained to fulfill the roles needed for this agrarian society. Education outside the farming community was limited to a small population of chosen people and selected professions, such as the Clergy (Miller, 2000b).
One of the main functions in life was to scan the immediate environment to find food and these basic skills had to be handed on by word of mouth. Survival of the fittest was the order of the day and it provided an essential platform for youth to acquire essential skills for life. This is characterized by the stone masons who were able to pass on the skills from generation to generation often with sons completing the work on a cathedral which their fathers had started. The Three R’s of Reading, Writing and Arithmetic form the foundation stones of all education and learners are expected to have a firm grounding in these basics in order to progress successfully through the grades.

The Second Dimension – Y Axis (Knowledge and Comprehension)

During the Industrial Era the economic power of a person, city or country came from the capacity to manufacture products. With the Industrial Revolution and vast numbers of inventions especially in the areas of mechanics, huge factories were erected everywhere to manufacture all kinds of products (Miller, 2000b). This era needed a class of citizen who was equipped to deal with the monotony of operating machinery resulting in a simplistic ‘factory model’ of education.

A characteristic of this dimension is rote learning based on a Pavlovian behavioristic model which required an education system to equip the masses with a fundamental education. The observable changes required that new behavioral patterns were repeated until it became automatic (Mergel, 1998). Society needed a cheap labour force to operate the machines for the capitalistic society. The ‘thinking’ was left to the people in power and education was simply a sausage factory where individuals were not encouraged to rise above the masses and where conformity was the order of the day. Teachers in such a system were viewed as gatekeepers and the source of all knowledge and where text books were regarded as all important.

Education is generally aimed at providing efficient and effective schooling for the masses who could perform a minimal set of skills on command in the factories. Creative thinking, problem solving skills and entrepreneurial endeavours were not required nor envisioned for an educational outcome in the industrial based economy and world picture (Miller, 2000c).

The Third Dimension – Z Axis (Spatial Orientation)

The third dimension raises an interesting possibility. If one had to compare the former two dimensions to travelling along a road, the curriculum would offer learners with choices at key stages in their development. The rest of the time would be spent following a predetermined, well-trodden paths which would lead to a possible reward system at the end of the road. It can be argued that many teachers are unable to make the transition to the third phase due to the nature of their own training as well as a lack of exposure to broader educational issues. Many teachers remain at the second level as they tend to operate at a two-dimensional level relying heavily on the ‘security blanket’ of the syllabus bound prescribed text book.

Adhering to the first two levels of the model can produce outstanding results but there is still the need to allow learners to expand their minds into the realms of creativity and innovation which may only be possible in the third, fourth and fifth dimensions. The transition to the third dimension also implies the ability to convey a richness to the educational journey that was not previously possible. By having the ability to fly as opposed to driving, learners are encouraged to transcend the two-dimensional mode of the text book. Such systems encourage a learner-centered approach by pushing the boundaries in which the students are able to safely discover far more of their own talents. The principal goal of education is to create men who are capable of doing new things, not simply of repeating what other generations have done (Burkhalter, McLean, & Jones, 2004:50).

The term learner-centered refers to environments that pays careful attention to the knowledge, skills, attitudes, and beliefs that learners bring to the educational setting. This term includes teaching practices that have been called "culturally responsive," "culturally appropriate," "culturally compatible," and "culturally relevant" (Ladson-Billings, 1995:465-491). The term also fits the concept of "diagnostic teaching" (Bell, O'Brien, & Shiu, Designing teaching in the light of research on understanding, 1980): attempting to discover what students think in relation to the problems on hand, discussing their misconceptions sensitively, and giving them situations to go on thinking about which will enable them to readjust their ideas (Bell, Diagnosing students' misconceptions, 1982:6-10) Teachers who are learner-
centered recognize the importance of building on the conceptual and cultural knowledge that students bring with them to the classroom (Bransford, Brown, & Cocking, 1999).

The third dimension favors a constructivist approach to learning based on the premise that learners all construct their own perspective of the world, through individual experiences and schemas. Constructivism focuses on preparing the learner to problem solve in ambiguous situations (Mergel, 1998). Taken a step further this includes the concept of constructionism in its simplest form only knows something if one can explain it (Yager, 1991:52-57).

Intimately connected with experience is a more educationally descriptive meaning of constructivism. Students come into a classroom with their own experiences and a cognitive structure based on those experiences. These preconceived structures are either valid, invalid or incomplete. The learner will reformulate his/her existing structures only if new information or experiences are connected to knowledge already in memory (Hanley, 1994). It is therefore imperative that educators must provide a richer learning environment for learners to attach meanings to new learning experiences.

Modern ICT intensive organizations expect employees to have a new set of digital skills. These skills are often not promoted in an information overloaded curriculum to the detriment of many prospective employees. On the other hand programmes like Formula One in Schools, Green Power and the 4X4 Challenge (4x4 Challenge Project, 2009) actively promote the use of ICTs to supplement the curriculum through extra mural activities.

The transition to the third level implies that the country as a whole is making or has made the transition to a knowledge-based economy. Organizations at the same time need to embrace this change and encourage the educators in the classroom to catalyse this reform process from a grass-roots level.

The dilemma facing outcomes based education is that employers of today are dissatisfied even with the high achieving graduates who have passed through the school system. Graduates tend not to be adequately prepared for the demands of today's more complex society which requires such competencies as creative thinking, group problem solving and decision making, as well as the capacity to learn more and more efficiently and effectively which is inherent in 21st century skills. This means that a more sophisticated view of knowledge and learning is required than the one held in the previous industrial era because the economy is now based on selecting, processing and applying information and creating new knowledge and applications (Miller, 2000a).

The Fourth Dimension – Time

Extending the learning process beyond the confines of the core curriculum and the walls of the classroom should be the goal of all educators to overcome the time constraints imposed by traditional classes and restrictive timetables. Many teachers have personal limitations which may inadvertently restrict the natural progress of many learners. By employing such educational tools as the internet, CD Roms, email and other modern technological teaching aids, these technologies may, if carefully planned, have the potential to circumvent some of these personal limitations. By applying creative methodologies the teacher can empower learners to internalize the learning process and embark on a journey of self-discovery well beyond the confines of the minimum standards imposed by the prescribed curriculum.

Exposure to other curricular material, searching for information to supplement lessons through collaborative projects and web publishing can transform dull lessons into stimulating educational encounters that may motivate students to extend themselves beyond what was achievable in a two-dimensional traditional teaching environment.

In such a four-dimensional environment the emphasis is on communications including both synchronous and asynchronous methods. Students are able to access the wisdom of experts around the world and receive an answer in a relatively short period of time. This feature can go a long way to allowing students to be become motivated to actively utilize the power of such things as Text Messaging, Blogs, Wikis, etc. Access to mobile learning devices and educational portals may suite the learning styles of many more learners in relation to face-to-face traditional teaching. This is embedded in the Web 2.0 phenomenon where learners are able to combine their collective intelligence through the use of social software (O'Reilly, 2005).
The Fifth Dimension – Global vision

It is necessary to highlight a fundamental difference between the use of the term ‘teacher’ and an ‘educator’ in this paper. Anyone can teach through a process of information transfer and reinforcement strategies but not everyone has the ability to turn a teaching situation into a life-long learning experience. Teachers acting as gatekeepers of knowledge are threatened by the perception that learners may know more than them. Educators on the other hand strive to release these learners to construct their own knowledge so that the educator can focus on those individuals who really need their attention.

It is suggested that the transition to the process of life-long learning can take place without students being exposed to all the facets of the other dimensions. However, it is suggested that educators should also make this transition themselves. By actively promoting the process of life-long learning, educators may be able to make a greater impact on the lives of the students that have been entrusted to their care. Adaptability is a key facet of this stage for without it teachers are doomed to remain in a two-dimensional text-book bound world.

At its most fundamental level, globalisation is about the monumental structural changes occurring in the processes of production and distribution in the global economy. These structural changes are responses by many global enterprises that confront tremendous pressures and fantastic opportunities presented by the increased application and integration of advanced information and communications technologies (ICTs) into their core business processes. Through the application of information and communications technologies, enterprises have the ability to diminish the impact of space, time and distance (Cogburn, 1998).

Our children live in a global, digital world—a world transformed by technology and human ingenuity. Given the rapid rate of change, the vast amount of information to be managed, and the influence of technology on life in general, students need to apply current skill sets, as well as develop new skill sets to cope and to thrive in this changing society (enGuage, 2003:3).

In order for learners entering the market place to be able to flourish in a global society they need to have acquired an essential set of skills which can only come about through exposure to the concepts of globalization. Traditional education has the teacher as a gatekeeper of information but the net generation is demanding not the raw information but rather the tools to sift through large volumes of data in order to synthesize new knowledge. This shift from broadcast to interactive is the cornerstone of the N-Generation (Net Generation). They want to be users-not just viewers or listeners (Tapscott, 1988:3).

A key characteristic of the fifth dimension also involves group work and a division of labour where learners are challenged to explore lateral thinking, creativity, problem solving and innovative types of challenges. It is essential that learners are provided with real-world tools and real-world scenarios such as those that are embedded in the following types of competitions that promote science, engineering and technology.

- Formula One in Schools (F1 in Schools, 2008)
- Green Power (Greenpower, 2008)
- 4x4 Challenge (4x4 Challenge, 2008)

Implications for Teacher Training

Too many teachers, too few educators

- As technologies penetrate the classroom environment and are embraced by net generation students, what is the new role for the teacher? Since the new media is drastically changing the learning milieu, so must the teachers change their roles and become more adapted to the new system of learning (Tapscott, 1988).
- There is a need to transform education from a teacher-centered behaviouristic to a learner-centered constructivist model by empowering the teachers to empower the learners through the integration of modern ICTs into all aspects of the learning process as depicted in Figure 3.
• There is an opportunity to develop a system to plot teachers and learners on a grid indicating their relative position in terms of this model of education. A prototype tool is being developed and tested for self-assessment to plan personal growth rather than for formal evaluation purposes. By indicating the relative position of the teachers, coupled to specific professional developmental strategies, it is conceivable that teachers will be able to manage their own upliftment programs. In this respect, the skills development process will have a meaningful context.

• Re-skilling of teachers is needed to bring the vast majority of them up to date regarding modern pedagogies involving ICTs. Collaboration, group work, project and theme-based learning, and the integration of ICTs into all subjects are but a few examples of the processes that need to be conveyed. It is recommended that policymakers acknowledge that the vast majority of teachers are becoming burdened with administrative matters and are unable to find the necessary quality time for personal professional growth. If they are not exposed to the benefits of new approaches to education, they will continue to practice teaching in the tried and trusted armchair approach and not actively encourage learners to develop essential 21st century skills.

• The development of management skills need to be given a higher priority. This includes the management of people and information with the view to knowledge sharing in an open learning environment.

• The concept of Learning Communities that are connected to a National Grid of Learning needs to be cultivated. Far too often teachers assume that they are the only resource when the classroom door is closed. The sharing of teaching and learning material can be a powerful method of uplifting teachers to the status of true educators or learning facilitators especially with the availability of on-line learning materials.

Learning Communities are not simply another educational fad or a modest type of school reform but an attempt to rebuild society's educational system on a post-modern cultural foundation that is democratic and person-centered rather than mechanical, as well as ecological and life-centered rather than driven exclusively by economic forces. We offer the idea of the "cooperative community lifelong learning center" as a seed for social and cultural renewal - a form of education that reclaims the organic qualities of learning from pre-mechanistic times for a post-modern culture in the making (Miller A., Differences between the 'Traditional' and 'Outcomes Based' Education, 2000a).

• Understanding the needs of the net generation especially as far as providing e-learning opportunities is concerned.

**Implications for Schools**

• Access to the internet is no longer a nice to have but an absolute necessity for all schools. A telecommunication infrastructure needs to be developed to provide connectivity for all schools as outlined in the EduNet Concept for South African schools. Provision is made for this in the South African Telecommunications Act 103 of 1996 and amended in 2001 (Department of Education, 2004). Until the real cost of access to the internet is brought down to an affordable level, the information gap is going to widen, placing more and more learners at an even greater disadvantage with no real means of deriving benefits from the technology revolution.
• Schools in South Africa are facing large class sizes as a fact of life. New methodologies coupled to access to modern technologies, can be employed to allow the more able students to circumvent the limitations that such a system usually places on the teacher. Given the appropriate training, it is conceivable that an innovative educator may be able to ‘reduce the class size’ to a more manageable size while enabling the more able students to meet the requirements of their subject by employing open learning strategies and sound classroom management.
• The Ulwazi Concept (Beyers, 2007) has demonstrated the potential of providing virtual interactive connected learning classrooms with the means to share educational resources including access to top educators who are in geographically separated schools.

Implications for Learners

• Education is what is left over when you forget all the facts that your teacher made you memorize when you were in school - Quote by Mark Twain. This implies that learners can and must take greater responsibility for their own learning. Leaving learners to their own devices is not being advocated. On the contrary. The educator must still be in control of the process, but their role is more a facilitator rather than physically dictating what happens at the learning interface where there is greater emphasis on self-discovery after the basic skills have been acquired.

Conclusions

Responding to the needs of the Net Generation is an ever increasing challenge especially where ICTs and other technologies are being introduced into the classrooms. Teachers are having to adapt to these changes through a process of upgrading their own skills to empower them to become better facilitators. By doing so that are able to unleash the innate potential of the learners entrusted to them.

The proposed five dimensional educational model may be perceived as a tool for such teachers to comprehend the complexity of the situations they may find themselves in. The tool is aimed at raising the awareness of the possible similarities and differences that may exist between the teacher and the net generation in their class. Armed with this information the teacher will be better prepared to make the transition from the two dimensional world of the textbook to be able to operate in the multi-dimensional world of that the youth operate in to be able to capitalize on what these learning environments have to offer.

References


Enhancing Eight Grade Students’ Scientific Conceptual Change and Scientific Reasoning through a Web-based Learning Program

Ya-Wen Liao and Hsiao-Ching She
Institute of Education, National Chiao Tung University, Taiwan, R.O.C. // linyin1001@yahoo.com.tw // hcshe@mail.nctu.edu.tw

ABSTRACT
This study reports the impacts of the Scientific Concept Construction and Reconstruction (SCCR) digital learning system on eighth grade students’ concept construction, conceptual change, and scientific reasoning involving the topic of “atoms”. A two-factorial experimental design was carried out to investigate the effects of the approach of instruction and students’ level of scientific reasoning on their pre-, post-, and retention-Atomic Achievement Test, Atomic Dependent Reasoning Test, and Scientific Reasoning Test. The control group (N=100) received conventional instruction whereas the experimental group (N=111) received an SCCR Web-based course. Results indicate that the experimental group significantly outperformed the conventional group on post- and retention-Atomic Achievement Test and Atomic Dependent Reasoning Test scores, and retention-Scientific Reasoning Test scores. Moreover, students with a higher level of scientific reasoning significantly performed better than students with a lower level of scientific reasoning, regardless of their scores on post- and retention-Atomic Achievement Test and Atomic Dependent Reasoning Test. This study successfully demonstrates that the experimental group students outperformed the conventional group students in the domains of concept construction, conceptual change and scientific reasoning. Moreover, students with a higher level of scientific reasoning were more able to successfully change their alternative conceptions.

Keywords
Conceptual change, scientific reasoning, Web-based learning

Introduction
The unique characteristics of the World Wide Web have led educational researchers and practitioners around the world to regard it as a potential tool for improving teaching and learning. All learning activities, except providing an authentic wet lab experience, seem possible on the Web. Accessing the internet and incorporating the World Wide Web into teaching activities are considered to be fundamental skills for science teachers in the twenty-first century (Didion, 1997).

Tuvi & Nachmias (2001), using taxonomy modified from that of Mioduser, Nachmias, Lahav, & Oren (2000), reviewed 93 Web sites that focused on introducing atomic structures. They found the following similarities: text was clearly the major means to present information; automatic/human and technical/content-based help were present in less than 18%; less than 6% had interactive images, animation or sound; none of those Web sites were inquiry-based; and memorizing was the main cognitive process.

This study clearly indicated that science teaching and learning theories are absent in most of the science Web-based learning programs, that pedagogical considerations are not sufficiently considered in the design of the Web-based learning program, and that the capability of the Web and multimedia is not fully used to enhance students’ science learning. Therefore, this study integrates both Dual Situated Learning Model (DSLM) and scientific reasoning theories to develop a Scientific Concept Construction and Reconstruction (SCCR) digital learning course to facilitate students’ science learning.

More importantly, there is still disagreement about whether Web-based learning or traditional teaching are more effective on students’ achievement and attitude. Many studies suggest no difference in test scores following Web-based and traditional courses. Although students gain more confidence with computer use in a Web-based course (Leasure, Davis, & Thievon, 2000), other studies find that students enrolled in a Web-based course perform worse in a final exam than students educated by conventional instruction (Wang & Newlin, 2000). However, others indicate an apparent increase in satisfaction on Web-based courses (Katz & Yablon, 2002). Therefore, one of the purposes of this study is to determine whether students receiving the SCCR Web-based course perform better than the conventional group in their atomic achievement scores. From the point of view of science educators, we claim that it
is very difficult to bring about conceptual change and scientific reasoning unless the instructional design is based on well-developed conceptual change theories and models. Thus, this study attempts to explore whether students who received SCCR Web-based instruction would outperform a conventionally educated group of students in their conceptual change and scientific reasoning.

Lawson & Weser (1990) did a study to measure pre-to post-instruction change among college students. They found that less skilled reasoners were initially more likely to hold a variety of nonscientific beliefs about life and were less likely to change some, though not all, of those beliefs. However, there is still a lack of study for specific well-designed instruction, including Web-based learning, comparing the performance of students with high and low levels of scientific reasoning in the area of concept construction and conceptual change. Therefore, this study intends to explore whether students with higher levels of scientific reasoning would be more likely to change their alternative conceptions than students with lower levels of scientific reasoning.

Conceptual Change Model: Dual Situated Learning Model (DSLM)

Conceptual change has been a major research area of science education over the past two decades (Duit & Treagust, 2003). Research on conceptual change is spurred by the great volume of findings from studies over the past thirty years on students’ alternative conceptions or pre-instructional conceptions about scientific concepts.

In view of the different perspectives of science education researchers (Brown, 1993; Clement, 1991, 1993; Posner, Strike, Hewson, & Gertzog, 1982; Steinberg & Clement, 1997) and of cognitive psychologists (Carey, 1985; Chi, Slotta, & de Leeuw, 1994; Thagard, 1992; Vosniadou & Brewer, 1987) on conceptual change, the author integrates the strengths from both sides to her theoretical construct and has developed the Dual Situated Learning Model (DSLM) (She, 2002, 2004b). The term “Situated learning” means that the design of conceptual change learning events needs to be based on the nature of science concepts and students’ beliefs of science concepts, in order to determine what essential mental sets are needed for constructing a more scientific view of the concepts. Moreover, each dual situated learning event needs to be situated on the prior dual situated learning event in order to help students connect different mental sets. The term “Dual” means that the learning events serve the following three dual functions: considering both the nature of science concepts and students’ beliefs of science concepts; creating dissonance with students’ pre-existing knowledge, and providing a new mental set for them to achieve a more scientific view of the concept; stimulating students’ motivation and challenging their beliefs of the concepts as well as challenging students’ ontological and epistemological beliefs of science concepts.

The author claims that conceptual change can not occur simply by creating cognitive conflict, but that conceptual change needs to provide students with new mental sets in addition to creating dissonance. Creating dissonance with students’ pre-existing knowledge can arouse students’ curiosity and interest, as well as challenging their epistemological and ontological beliefs about scientific concepts. Motivation is embodied in the design and the learning process of a dual situated learning event, which requires students to actively engage in the event prediction, visualize what actually happens and explain why it is different from their prediction, thereby stimulating their curiosity and interest. Providing a new mental set should be the platform upon which knowledge reconstruction can occur. Students must comprehend and believe the new mental set in order for conceptual change to happen. This can be fostered by any type of instructional activity, such as analogy, modeling, discrepant events, and inquiry activities, as long as it provides students with opportunities to visualize what actually happens in order to reconstruct new mental sets. DSLM emphasizes that students should be actively engaged in the conceptual change learning process. All of the preceding criteria should serve as the basis for developing each dual-situated learning event in order for it to be a success. Moreover, the author claims that conceptual change can not happen due only to a discrepant event; the changes need a series of events to increase the success of students’ conceptual change (She, 2002). The number of dual situated learning events required would depend on the number of mental sets that students lack for constructing a more scientific view of the concepts. More importantly, each dual situated learning event should be connected with the others and needs to build upon a prior dual situated learning event.

In order to implement this theory, the DSLM has been transformed into a six-stage learning/instructional model.

Stage 1: Examining the attributes of the science concept. This stage provides information about which essential mental sets are needed in order to construct a scientific view of the concept.
Stage 2: Probing students’ alternative science concept, which requires understanding the students’ beliefs concerning the science concept.

Stage 3: Analyzing which mental sets students lack for the construction of a more scientific view of the concept.

Stage 4: Designing dual situated learning events, according to the stage 3 results.

Stage 5: Instructing with dual situated learning events, giving students an opportunity to make predictions, provide explanations, confront dissonance, and construct a more scientific view of the concepts.

Stage 6: Instructing with a challenging dual situated learning event. It provides an opportunity for the students to apply the mental sets they have acquired to a new situation in order to ensure a successful conceptual change.

Application of the DSLM instructional approach has evidenced effective conceptual change in middle school students in the topics of air pressure and buoyancy (She, 2002), thermal expansion (She, 2003), heat transfer (She, 2004b), dissolution and diffusion (She 2004a), as well as meiosis and mitosis (Tang, She, & Lee, 2005). This study attempts to use the DSLM to design an SCCR digital learning program.

Scientific Reasoning

Lawson (2003) indicates that the view of development of advanced reasoning and its emphasis on alternative conceptions leads quite naturally to a theory of science instruction which has the potential to unite two previously competing research traditions in science education. In that study, Lawson discussed the relationship between students’ alternative concepts and reasoning ability. Students need to be aware of their own alternative conceptions and scientific conceptions, together with the evidence and reasoning that bears on the validity of the alternative conceptions, in order to modify their alternative concepts. In other words, they must be able logically to see how the evidence supports the scientific conceptions and contradicts the prior alternatives. Park & Han (2002) assumed that deductive reasoning plays an important role in conceptual change. However, due to various unknown reasons, reasoning is not always activated and used spontaneously in students’ minds in the process of conceptual change. The conceptual change model is based on reasoning and the design of each dual situated learning event requires that students provide explanations. However, DSLM may not be efficient in activating the spontaneous use of scientific reasoning, by students, to cope with conceptual change. Thus this study combines both scientific reasoning and DSLM theories in designing the SCCR digital learning program.

Lawson & Worsnop (1992) again found that the more skilled reasoners in a sample of high school students were less likely to hold pre-instruction misconceptions regarding evolution and special creation. Lawson & Weser (1990) did a further study to measure pre- to post-instruction change among college students and found that less skilled reasoners were initially more likely to hold a variety of nonscientific beliefs about life and were less likely to change some, though not all, of those beliefs. Therefore, one of the purposes of this study is to explore whether students with a higher level of scientific reasoning would be more likely to change their alternative conceptions and hold more correct scientific conceptions.

Previous Research on Students’ Alternative Conceptions Involving Atoms

Garnett, Garnett, & Hachling (1995) review previous studies on students’ alternative conceptions for the particulate nature of matter, and specifically point out that some students consider atoms to be alive. Harrison & Treagust (1996) report that high school students view atoms as either visible under a microscope or too small to be seen at all, are not sure whether all substances contain atoms, consider that atoms are alive like cells and can grow and divide, view atoms as balls or spheres but are not sure about the components inside atoms, view the texture of atoms as being most like a hard polystyrene sphere, are not aware of the electron shell or electron clouds, and have a consistent image of atoms as being protected by an outer shell. Lee, Eichinger, Anderson, Berkheimer, & Blakeslee (1993) find similar results about the size of molecules, so that even after instruction, students still believe they can see molecules with a microscope. Lee et al. (1993) show that students believe that atoms are hard. Griffiths & Peterson (1992) report that students tend to believe that all atoms weigh the same, atoms are ball-shaped, there is nothing inside atoms, something exists between atoms, and electrons move within the shell.
Harrison & Treagust (2001) further reported that high school students perceive that there is a large distance between electrons and nucleus and that an atom is mostly space with the majority of its mass in the nucleus. As well, high school students are not able to differentiate between electron shells and electron clouds. Taber (2001) reported that students often have little difficulty in explaining how size increases down the periodic table, as more electron shells are added. The design of the atomic related Web-based learning events in this study considers some alternative conceptions described above to be one of the major sources.

**Purpose**

This study reports the development of the Scientific Concept Construction and Reconstruction (SCCR) digital learning system that was developed based on the theories of the Dual Situated Learning Model (DSLM) and scientific reasoning, and further examines its impact on eighth grade students’ concept construction, conceptual change and scientific reasoning involving the concept of atoms. A two-factorial experimental design was carried out to investigate the effects of the approaches of instruction and the students’ levels of scientific reasoning on their pre-, post-, and retention-Atomic Achievement Test (AAT), Atomic Dependent Reasoning Test (ADRT), and scientific reasoning Test (SRT). The control group (N=100) received conventional instruction, whereas the experimental group (N=111) received the SCCR digital learning course. Students were classified into high and low levels of scientific reasoning ability according to their Scientific Reasoning Test (SRT) scores.

Three goals serve as the guideline for examining the impact of the SCCR. The first is to document that students receiving an SCCR Web-based course perform better than conventional group in their atomic conceptual understanding achievement. The second is to demonstrate that SCCR Web-based instruction can be more effective than conventional instruction in facilitating the experimental group students’ conceptual change as well as scientific reasoning in the “atoms” unit. The third is to confirm that students with higher levels of scientific reasoning ability would be more able to overcome their alternative conceptions than the students with lower levels of scientific reasoning.

![Figure 1. Web page of teacher authoring module](image-url)
Development of SCCR Digital Learning Project

The SCCR project was developed based upon the theories of DSLM and scientific reasoning as the foundation for developing the Web-based learning content. The current SCCR digital learning project includes several units: the electricity unit and buoyancy unit in physics, the atom unit and the combustion unit in chemistry, and the genetics unit in biology. All of the units have been implemented in several middle schools’ science classes and received similar encouraging results. This paper reports an impact of SCCR on students’ learning from one the units in chemistry: atom.

SCCR Digital Learning System Software and Operating Requirement

The SCCR platform was FreeBSD running with an Apache WWW server. The core of the SCCR system was programmed in PHP and Java Applet and JavaScript and it worked with MySQL to efficiently handle extremely large data sets and analytical programs. By using PHP, we can receive user input and process it on the server side, and then dynamically generate the next new pages. We also used the free software Simple Machine Forum (SMF) which is programmed in PHP to build up a discussion forum. To synchronize the data between SCCR and SMF, we rewrote some codes in SMF.

For the authoring interface, there were three main pages (Normal, Advance and Special) and ten different kinds of functional pages (eg. junction, comparisons, correct answer, keyword to multipath, etc.), depending on the nature of the science concept involved and students’ learning progress (Figure 1).

![Figure 2. Students answer and reasoning before and after the web-based dual situated learning events](image)

Characteristics of the SCCR:

*Facilitate students' conceptual change and scientific reasoning*

In order to facilitate students’ conceptual change and scientific reasoning, the design of each Web-based dual situated learning event requires students to provide an answer to the driving question, along with their reasons, before proceeding with various activities. In addition to writing down their reasons, students were required to choose the best reasons for what they had written down. The system generated an appropriate HTML page for each
individual student according to the answers and reasons selected before each event. The same driving question was asked again after the event to examine the students’ conceptual change and their reasoning changes. In order to activate students’ use of scientific reasoning, we specifically restructured the process. Students could view their answers and their written reasons before and after events. They could also view the correct answers. We then requested that students justify why they changed or stayed with their original thoughts after the learning events. This process was purposely intended to direct students’ minds to spontaneously become involved in the process of scientific reasoning and conceptual change (figure 2).

Facilitate individualized science learning

A Web-based learning system must not only allow students to navigate between different pages but must also help students better achieve their learning goals. It is also necessary to provide more sophisticated mechanisms that modify the navigation alternatives by a procedure for individual student adaptation. Adaptive learning can offer flexible solutions that dynamically adapt content to fit individual real-time learning needs. Very often, students hold different alternative conceptions or different reasoning patterns. Adaptive learning is therefore included in the design of SCCR to provide students with different learning paths to maximize their science learning.

Adaptive learning to individual student needs is implemented by presenting students with different HTML pages, depending on their prior alternative concepts and their previous reasoning results. The adaptive linking is used to modify the link structure depending on the state of students’ conceptions and reasoning. This leads to the development of an individual course, with a dynamically changing link structure, depending on each individual student’s progress.

Different learning events are provided depending on what alternative concepts and reasoning patterns the students already possesses. Some of the units of SCCR are specifically designed to provide students with different learning events according to the choices of answers and reasons they made regarding either accuracy of concepts and reasoning, and thus further determine which learning pathway the students need to work on.

Figure 3. The 2D-atomic structure model

Design of SCCR content: Unit for Atoms

The design of the Web-based program for the unit on atoms specifically focuses on those students’ alternative concepts of atoms, as found in previous studies (Griffiths & Peterson, 1992; Harrison & Treagust, 1996; Lee et al., 1993). It is clear that students’ alternative concepts of atoms are due to the invisible molecular and dynamic nature of
atoms, so students have difficulty constructing concepts related to atoms. Therefore, the design of this unit on atoms provides students with 2D- and 3D-models to help them develop more scientific views of atomic structure. For instance, students can interact with one of the models. They can then visualize the motion of, and number of electrons in, each shell, or the number of protons and neutrons at the nucleus of the atom (Figure 3). As another example, with the model of chemical reaction, students can see how atoms react with other atoms at the microscopic level.

The design of the unit on atoms is based upon the six stages of DSLM described above. The mental sets which students lacked were categorized into four topics to further develop the following series of four dual situated learning events:

1. Elements and compounds: identify the difference between compounds and elements, and understand how compounds can be separated into elements through chemical means.
2. Atomic structure: concerning the components of atoms and their properties as well as the motion of electrons.
3. Arrangement of electrons and chemical reactivity of elements: understand that the arrangement of electrons in a noble gas is stable because there are eight electrons on the outer shell, and the alkali metals such as sodium and potassium have greater chemical reactivity because there is only one electron at the outer shell of the atom.
4. Atomic and chemical reactions: the nature and number of electron transfers occurring when two different atoms react with each other to form a compound, as well as the relationship between the donating/receiving electrons and the electron arrangement of atoms during a chemical reaction.

**Subjects and Procedures**

A total of 211 eighth graders participated in this study, recruited from six average-achievement classes of a middle school in Taiwan. One hundred and eleven of the participants received the Web-based DSLM learning (the experimental group) in three classes and the remaining 100 students received conventional lecture-based instruction (the control group) in the other three classes. Based on the students’ scientific reasoning pre-test scores (pre-SRT), the students were further classified as transitional (higher level of scientific reasoning) (N=69) and concrete operational level (lower level of scientific reasoning) (N=142) students. A pretest for AAT was administered to all students before the instruction. The independent t-test on the pretest scores of atomic related concepts showed that the two groups had an equivalent initial ability (t = 1.5, p = .12). All of the students received the AAT, ADRT, and SRT before instruction, after instruction and eight weeks after the instruction.

The mandated class periods for physical science are four class periods per week. The school schedule allocated approximately 12 class periods on the atom unit which included two class periods for laboratory activities. There were a total of ten class periods over four weeks for learning the atom unit, with each class period lasting about 45 minutes. Our study took place over two class periods in the first week of the four-weeks learning of the atom unit, then three class periods in the second week and third week, and two class periods in the fourth week. Both groups of students did the same experiments in their respective laboratory periods.

Students who received the SCCR Web-based learning program learned and interacted with the Web courses individually, so they learned at their own pace. Teachers served as facilitators to provide help if students requested it and the researcher also served as an assistant to help with any problems that arose from using the Web course.

For the conventional group, teachers mainly gave lectures and used the textbook as the primary major method of instruction. Students were mainly passive when receiving knowledge from the teacher and the textbook. The teachers also arranged discussions and did the same experiments in the laboratory for students.

** Instruments**

**Atomic Achievement Test (AAT)**

The AAT is a multiple choice diagnostic instrument developed for this study to measure students’ concepts of atom-related concepts before, directly after, and three months after receiving the atom unit of the SCCR digital learning program (Appendix 2). The content validity was established by a panel of eight evaluators, ensuring that the items
were properly constructed and relevant to the atom unit Web-based learning materials we developed. The questions requiring students to use their deeper information processing ability are mainly concerned with the analysis and synthesis. There are nine items for topic 1 and the other three topics each have eight items, for a total of 33 items. Students receive one point for each question they answered correctly, so the highest score is 33. The Cronbach $\alpha$ of the Atomic Achievement Test (AAT) was 0.74 for the pretest, 0.90 for the posttest, and 0.93 for the retention-test.

Atomic Dependent Reasoning Test (ADRT)

The ADRT is a two-tier multiple choice diagnostic instrument that was developed for this study to measure the degree of students’ conceptual change involving atom-related concepts before, directly after, and three months after receiving the atom unit of the SCCR digital learning program (Appendix 3). The ADRT requires students to use scientific reasoning to reason and process atom-related concepts in order to answer both tiers correctly. The content validity was established by the same panel of eight evaluators, ensuring that the items were properly constructed and relevant to the atom unit Web-based learning materials we developed. There are eight items for each topic, and each item contains two tiers, with the first tier to check whether students have scientific concepts/or alternative concepts, and the second tier to require students to use scientific reasoning on atom-related concepts. There are 32 items and each item has two tiers. Students need to answer both tiers correctly in order to receive one point, so the highest score is 32. The Cronbach $\alpha$ of ADRT was 0.88 for the pretest, 0.95 for the posttest, and 0.95 for the retention-test.

Scientific Reasoning Test (SRT)

The SRT is a two-tier multiple-choice diagnostic instrument that was developed to measure students’ scientific reasoning before, directly after, and three months after receiving the atom unit of the SCCR digital learning program (Appendix 1). The SRT was originally developed by Lawson in 1978 and was modified in 2003. It measures students’ conservation, proportional thinking, identification and control of variables, probabilistic thinking, correlative thinking and hypothetic–deductive ability. There are 12 items, and each item contains two tiers—the first tier for choosing the answer, and second tier for using the thinking ability mentioned above. Students need to answer both tiers correctly in order to receive one point, so the highest score is 12. The Cronbach $\alpha$ of the Scientific Reasoning Test (SRT) was 0.78 which is close to Lawson’s result for the Cronbach $\alpha$.

Results

Atomic Achievement Test (AAT)

Two-factor MANCOVA was conducted to examine the effects of instructional approaches and scientific reasoning levels using post- and retention-AAT scores as the dependent measures, and students’ pre-AAT scores as the covariate. Table 1 summarizes the results of the two-factor MANCOVA: specifically, instructional approaches (Wilk’s $\Lambda=0.91$, $p= 0.000$) and students’ level of scientific reasoning (Wilk’s $\Lambda=0.95$, $p= 0.007$) both have statistically significant effect on the performance of post- and retention-AAT. There is no interaction found between instructional approaches and students’ level of scientific reasoning. In short, both the instructional approaches and the level of scientific reasoning have significant impact on students’ performance of atomic achievement.

<table>
<thead>
<tr>
<th>Table 1. Multivariate Analysis of Covariance (MANCOVA) of Post- and Retention- of Atomic Achievement Test (AAT) Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of Variance</td>
</tr>
<tr>
<td>Covariates</td>
</tr>
<tr>
<td>Pre-test scores</td>
</tr>
<tr>
<td>Group memberships</td>
</tr>
<tr>
<td>Instructional approaches</td>
</tr>
<tr>
<td>Levels of Scientific Reasoning</td>
</tr>
</tbody>
</table>

Note. **$p<0.0001$, *$p<0.001, \ p<0.01$**

Therefore, the following main effect for the instructional mode and level of scientific reasoning was performed.
First, one-factor MANCOVA was performed to examine the effect of instructional approach factor on both post-test and retention-test, as shown in Table 2. One-factor MANCOVA shows a significant effect for instructional approaches (Wilk’s $\Lambda=0.85$, $p=0.000$). Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approaches on post- and retention-AAT. This indicated that the effects for instructional approaches on both post-AAT ($F=12.87$, $p=0.000$) and retention-AAT ($F=38.77$, $p=0.000$) were significant. Thus, the students’ post- and retention-AAT were significantly affected by the instructional approach. The post-hoc analysis for the main effect suggests that the Web-based instructional group performed significantly better than the conventional group (Web-based instruction > conventional instruction, $p_{\text{post}}=0.000$, $p_{\text{retention}}=0.000$) on post- and retention-AAT. In summary, the Web-based group outperformed the conventional group on both post- and retention-performance of atomic achievement.

<table>
<thead>
<tr>
<th>Instructional Approaches</th>
<th>Wilk’s $\Lambda$</th>
<th>Univariate F Post-test</th>
<th>Univariate F Retention-test</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.85***</td>
<td>12.87***</td>
<td>38.77***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. MANCOVA and ANCOVA of Instructional Approaches and 2-Levels of Scientific Reasoning of Post- and Retention- of Atomic Achievement Test (AAT) Scores

Table 3 summarizes the results of the two-factor MANCOVA: specifically, instructional approaches (Wilk’s $\Lambda=0.75$, $p=0.000$) and students’ levels of scientific reasoning (Wilk’s $\Lambda=0.92$, $p=0.000$) both have statistically significant effect on the performance of post- and retention-ADRT. There is no interaction between instructional approaches and students’ level of scientific reasoning. In short, both instructional approaches and level of scientific reasoning have significant impact on students’ conceptual change. Therefore, the following main effects for the instructional approaches and level of scientific reasoning were performed separately.

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Wilk’s $\Lambda$</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Multivariate F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test scores</td>
<td>0.84</td>
<td>2</td>
<td>199</td>
<td>19.09***</td>
</tr>
<tr>
<td>Group memberships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional approaches</td>
<td>0.75</td>
<td>2</td>
<td>199</td>
<td>33.92***</td>
</tr>
<tr>
<td>Levels of Scientific Reasoning</td>
<td>0.92</td>
<td>2</td>
<td>199</td>
<td>8.37***</td>
</tr>
</tbody>
</table>

Note: ***$p<0.0001$, **$p<0.001$, *$p<0.01$
First, one-factor MANCOVA was performed to examine the effect of instructional approach factor on both post- and retention-ADRT. One-factor MANCOVA shows a significant effect for instructional approaches (Wilk’s Λ=0.70, \( p=0.000 \)). Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approaches on post- and retention-ADRT. It indicated that the effects for instructional approaches on both post-ADRT (\( F=82.06, p=0.000 \)) and retention-ADRT (\( F=47.68, p=0.000 \)) were significant. Thus, the students’ post- and retention-ADRT were significantly affected by the instructional approach. The post-hoc analysis for the main effect suggests that the Web-based instructional group performed significantly better than the conventional group (Web-based instruction>conventional instruction, \( p_{(post)}=0.000, p_{(retention)}=0.000 \)) on both the post-test and the retention-test. This shows that the Web-based group outperformed the conventional group on the post- and retention-performance of conceptual change.

Second, the same procedures were performed for the effect of the level of scientific reasoning. One-factor MANCOVA shows a significant effect for the level of scientific reasoning (Wilk’s Λ=0.89, \( p=0.000 \)). Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the level of scientific reasoning on post- and retention-ADRT. It indicated that there was a significant effect for the level of scientific reasoning on both post-ADRT scores (\( F=22.78, p=0.000 \)) and retention-ADRT (\( F=13.32, p=0.000 \)). Thus, the students’ post- and retention-ADRT were significantly affected by the level of the scientific reasoning. The post-hoc analysis for the main effect suggests that the transitional level of students in the scientific reasoning group performed significantly better than the students in the concrete operational group (transitional >concrete operational, \( p_{(post)}=0.000, p_{(retention)}=0.000 \)) on post- and retention-ADRT. This indicated that students with higher levels of scientific reasoning outperformed the conventional group on their post- and retention-performance of conceptual change.

### Scientific Reasoning Test (SRT)

Two-factor MANCOVA was conducted to examine the effects of instructional approaches and scientific reasoning levels using post- and retention-SRT scores as the dependent measures, and students’ pre-SRT scores as the covariate. Table 4 summarizes the results of the two-factor MANCOVA, showing that only instructional approach has a statistically significant effect on the performance of post- and retention-SRT (Wilk’s Λ=0.95, \( p=0.006 \)). There is no interaction between instructional approach and students’ level of scientific reasoning. In short, only instructional approach has a significant impact on students’ scientific reasoning performance. Therefore, the following main effect for the instructional approaches was performed.

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Wilk’s Λ</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Multivariate F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test scores</td>
<td>0.79</td>
<td>2</td>
<td>205</td>
<td>27.90 **</td>
</tr>
<tr>
<td>Group memberships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional approaches</td>
<td>0.95</td>
<td>2</td>
<td>205</td>
<td>5.29 *</td>
</tr>
<tr>
<td>Levels of Scientific Reasoning</td>
<td>0.99</td>
<td>2</td>
<td>205</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Note. **p<0.001, *p<0.01, p<0.01**

One-factor MANCOVA was performed to examine the effect of instructional approach factor on both post- and retention-SRT. One-factor MANCOVA shows a significant effect for instructional approaches (Wilk’s Λ=0.92, \( p=0.000 \)). Then univariate F (one-factor ANCOVA) was performed to independently examine the effect of the instructional approaches on post- and retention-SRT. It indicates that the effect for instructional approaches on retention-SRT (\( F=16.06, p=0.000 \)) was significant. The post-hoc analysis for the main effect suggests that the Web-based instructional group performed significantly better than the conventional group (Web-based instruction>conventional instruction, \( p_{(retention)}=0.000 \)) on retention-SRT. In short, the Web-based group outperformed the conventional group in the retention-performance of scientific reasoning performance.
Discussion and Conclusions

This study reports a Web-based learning project that was developed based on the Dual Situated Learning Model (DSLM) and scientific reasoning theories in order to promote middle school students’ conceptual change and scientific reasoning ability. This study is a major step from previous Web-based instructional learning programs, as it brings well-developed science learning and pedagogy theories and models into an SCCR Web-based learning program. In addition, personal adaptivity is implemented into the design of SCCR to provide individual students with different HTML pages, depending on each student’s prior alternative conceptions and scientific reasoning results. Moreover, SCCR also develops many unique functions (eg. junction, comparisons, correct answers, keyword to multipath, etc.) for teachers to develop the Web-based learning content to make learning more efficient.

The results of this study show that the students who received a SCCR Web-based learning course significantly outperformed the conventional group of students in both post- and retention-AAT. Many studies suggest that there is no difference in test scores between Web-based and traditional format courses, although students may gain more confidence with computers in a Web-based course (Leasure et al., 2000); other studies find that students in a Web-based course actually perform worse than conventional students in final exams (Wang & Newlin, 2000); and still others indicate an apparent increase in satisfaction from Web-based courses (Katz & Yablon, 2002). The result of AAT adds a positive documentation to the current research that students who receive a Web-based learning course can perform better than a conventional group’s students in their atomic concept construction, especially when the Web-based learning course is effectively designed.

With the ADRT, it shows that students who received SCCR Web-based learning significantly outperformed the conventional group of students in both post- and retention-ADRT. This study definitely demonstrates that combining both the DSLM conceptual change model and scientific reasoning into a Web-based learning program does facilitate the experimental group of students’ conceptual changes involving atomic concepts. Our study demonstrated that students’ conceptual change can be facilitated through a well-designed conceptual change model and scientific reasoning theories.

With the SRT, students who received Web-based learning statistically significant outperform the conventional group of students on retention-SRT. It is possible that since there were only ten class periods over four weeks of instruction, this length of time might be too short to make a very large difference between the two groups in their immediate test. However, it supports the contention that students’ performance of scientific reasoning can be retained after eight weeks of learning from a well designed learning program.

The possible explanation of AAT and ADRT not showing any interaction between the instructional approaches and level of scientific reasoning is that students with a higher level of scientific reasoning perform better on the concept test and conceptual change test than students with a lower level of scientific reasoning no matter which types of instruction they received. It is especially noteworthy that students who had a higher level of scientific reasoning (transitional) performed significantly better than those with a lower level of scientific reasoning (concrete operational) on both post- and retention-AAT. This implies that students with higher levels of scientific reasoning would outperform those with lower levels of scientific reasoning on their atomic concepts construction. Moreover, students with a higher level of scientific reasoning (transitional) performed significantly better than students with a lower level (concrete operational) on both post- and retention-ADRT. Lawson & Weser (1990) conducted a study to measure pre- to post-instruction change among college students. They found that less skilled reasoners were initially more likely to hold a variety of nonscientific beliefs about life and were less likely to change some. This supports our findings. Our result confirms the findings of Lawson & Weser (1990), and further demonstrates that students with higher levels of scientific reasoning ability are more able to overcome their alternative conceptions and make the conceptual change more effectively.

To summarize, this study clearly demonstrates that uniting DSLM and the scientific reasoning theory in a Web-based learning environment promotes the experimental group of students’ atomic concept construction as well as their conceptual change more efficiently when compared with the group of conventionally-instructed students. This indeed provides evidence that using well developed theories and models of conceptual change and scientific reasoning as the basis for developing Web-based instructional materials and learning environment makes a significant difference. In addition, this study further contributes to the field of study by demonstrating that students
with higher levels of scientific reasoning would be more likely to evidence conceptual change and construct more correct scientific conceptions as compared to students with lower levels of scientific reasoning.

The implications are provided by the present study in the following:
1) It is highly suggested that Web-based learning can be successful if the design of the Web-based course is based on a specific learning or pedagogy theory.
2) Students’ successful learning or conceptual change is possible if the design of Web-based learning is based on well-developed learning or pedagogy theories.
3) More empirical studies are needed to confirm the results that students with higher levels of scientific reasoning are more likely to change their alternative conceptions, as well as confirming that specific learning or pedagogy theory-based Web learning can effectively promote science learning.

Acknowledgements

This research is based on work supported and funded by the National Science Council (NSC 95-2511-S-009-002), Taiwan, R.O.C.

References


Disabilities and e-Learning Problems and Solutions: An Exploratory Study

Catherine S. Fichten¹, Vittoria Ferraro², Jennison V. Asuncion², Caroline Chwojka², Maria Barile², Mai N. Nguyen², Ryan Klomp³ and Joan Wolforth⁴

¹Psychology Department and Adaptitech Research Network, Dawson College, Montreal, Canada // catherine.fichten@mcgill.ca
²Adaptitech Research Network, Dawson College, Montreal, Canada // vferraro@dawsoncollege.qc.ca // asuncion@alcor.concordia.ca // caroline_chwojka@hotmail.com // mbarile@dawsoncollege.qc.ca // blitzball_game@hotmail.com
³299 Somerset St. W. # 308 Ottawa, Canada // ryanklomp@rogers.com
⁴Office for Students with Disabilities, McGill University, Montreal, Canada // joan.wolforth@staff.mcgill.ca

ABSTRACT
This study explored e-learning problems and solutions reported by 223 students with disabilities, 58 campus disability service providers, 28 professors, and 33 e-learning professionals from Canadian colleges and universities. All four groups indicated, via online questionnaires, problems with: accessibility of websites and course/learning management systems (CMS); accessibility of digital audio and video; inflexible time limits built into online exams; PowerPoint/data projection during lectures; course materials in PDF, and lack of needed adaptive technologies. Students also mentioned technical difficulties using e-learning and connecting to websites and CMS, problems downloading and opening files, web pages that would not load, video clips taking too long to download, poor use of e-learning by professors and their own lack of knowledge working with e-learning. Disability service providers, too, mentioned the poor use of e-learning by professors as well as poor accessibility of course notes and materials in many formats. E-learning professionals noted difficulties with inaccessible course notes and materials. Professors identified mainly problems raised by the other groups. Sixty-seven percent of students, 53% of service providers, 36% of e-learning professionals and 35% of professors indicated that at least one of their three e-learning problems remained unresolved. We discuss how the different roles and perspectives of the four participant groups influence their views, and make recommendations addressing identified common e-learning problems.

Keywords
Disability, College students, Professors, E-learning, Accessibility, Problems and solutions

The increased use of information and communication technologies (ICTs) in most sectors of society and recent developments in adaptive hardware and software have allowed individuals with disabilities to do things that were difficult or impossible for them to do in the past. For example, it has allowed people who are blind to read using text-to-speech technology, people who are deaf to communicate using chat programs, and people with difficulties using their hands or arms to write and communicate using dictation software (Fichten, Asuncion, Barile, Fossey, & De Simone, 2000). In the postsecondary arena, e-learning (i.e., ICTs used in teaching) are extensively used by professors in virtually all colleges and universities, not only in the United States, but also in Canada (Abrami et al., 2006) and the United Kingdom (Weller, Pegler, & Mason, 2005). Whether it is PowerPoint presentations in class, the use of web-based discussions to further in-class conversation, or courses delivered completely over the Internet, it is clear that such technologies used by professors are here to stay.

But how well do the ICTs used by professors in teaching postsecondary courses (i.e., e-learning) meet the needs of students with different disabilities? How successfully do these ICTs interact with adaptive hardware and software that some students with disabilities require? How accessible is the growing array of available e-learning on campus (Konur, 2007; Waddell, 2007)? These are important questions because the numbers of students with disabilities in postsecondary education have been rising both in Canada (Fichten, Jorgensen, Havel, & Barile, 2006; Tremblay & Le May, 2005) and the US (National Council on Disability, 2003), where a recent large scale study showed that in 2003–2004, 11% of undergraduates had a disability (Snyder & Dillow, 2007). In addition, during the past few years, skill in using ICTs has become mandatory in postsecondary education and the workplace (e.g., Ezziaine, 2007; Stodden, Conway, & Chang, 2003). For example, a recent investigation shows that computer use on the job is associated with higher salaries for employees both with and without disabilities (Canadian Council on Social Development, 2004), and that for people with disabilities this is especially important (e.g., Kruse, Krueger, & Drastal, 1996). Abrami et al. (2006), who recently showed how important e-learning initiatives are in Canadian postsecondary education, also noted that we know very little about the e-learning needs and concerns of students with disabilities.
E-learning can promote the inclusion of students with various disabilities (Di Iorio, Feliziani, Mirri, Salomoni, & Vitali, 2006). For example, online courses provide enhanced opportunities for people who, because of climate, health, transportation or physical accessibility, experience barriers to attending classroom-based courses (e.g., Debenham, 2002). Similarly, in traditional classes students who have print impairments can access course notes and handouts on the course website without assistance, so long as these are designed to be accessible.

In spite of the tremendous opportunities afforded by e-learning for learners with disabilities there are a variety of barriers that interfere with their effective use. A brief primer on how students with different disabilities use ICTs can be found in Fichten, Asuncion, Barile, Fossey, and De Simone (2000). A key concern is that faculty and individuals accountable for supporting and implementing e-learning within postsecondary institutions, in the rush to integrate technology into teaching, fail to think about the accessibility needs of students with various disabilities (Bissonnette, 2006). For example, those in charge of supporting and deploying e-learning generally do not confirm ahead of time whether newly purchased academic software is compatible with adaptive software that reads what is on the screen to individuals with print impairments. PowerPoint presentations in class, if not posted online ahead of time, can cause difficulties for students with visual and other disabilities requiring adaptive software to read and follow the presentation. Video clips posted on a course website with no captioning can pose problems for students with hearing impairments. Websites can pose problems for students with learning, visual, and neuromuscular disabilities even when they use adaptive software such as screen magnification, screen reading, and dictation software (Burgstahler, Corrigan, & McCarter, 2005; Bohman, 2007; Roberts & Stodden, 2005; Sharpe, Johnson, Izzo, & Murray, 2005).

At least four postsecondary groupings have a stake in accessibility and e-learning in colleges and universities: the students themselves, service providers who provide disability related services to the campus community, professors who use and implement e-learning in their courses, and the e-learning professionals on campus who provide leadership, choose e-learning for campus-wide use and provide help and assistance with e-learning and other ICTs. Because of their different perspectives, these four groups are likely to have different views about e-learning accessibility (Fichten, Asuncion, Barile, Fossey, & Robillard, 2001).

There have been studies related to the ICT needs and concerns of students with disabilities, including e-learning, where the participants were campus disability service providers (e.g., Stodden, Roberts, Picklesimer, Jackson, & Chang, 2006), postsecondary graduates (Sharpe et al., 2005), assistive technologists (Thompson, 2004), and professors (Vogel, Leyser, Burgstahler, Sligar, & Zecker, 2006). These studies all had different goals, used different questions, and had different sampling limitations, making their findings impossible to compare. In addition, in none of these studies did a number of different groups answer the same questions about the problems they experienced with e-learning and how these were resolved. Perhaps more important, none of these studies included the end-users — current students with disabilities — in their samples.

In the present study, which forms part of a larger investigation of e-learning (Asuncion, Fichten, & Barile, 2007), we examine the perspectives of the four key stakeholders about the accessibility of e-learning: postsecondary students with a variety of disabilities, campus disability service providers, professors, and e-learning professionals. We examine both problems and solutions as experienced in Canadian junior/community colleges and universities and also assess the benefits of e-learning as experienced by the students themselves. Based on the findings, we make recommendations about addressing common e-learning problems encountered in postsecondary education and about how the different roles and perspectives of the four participant groups influence their views. It is important to note this is an exploratory, descriptive study that is not theoretically based. Its main objective is to compare the views of the four groups, to suggest hypotheses for future investigations, and to propose recommendations based on available information.

**Method**

**Participants**

Two hundred twenty-three Canadian university and junior/community college students with various disabilities (74 males, 149 females), 58 campus disability service providers (15 males, 43 females), 28 professors (10 males, 18 females), and 33 campus e-learning professionals (16 males, 17 females) participated. Participants studied or worked at 318 Canadian postsecondary institutions.
Students with disabilities had taken at least one course in the last three years that used some form of e-learning. Eighty-two students were enrolled in a Canadian junior/community college and 136 in a university; 5 did not specify this information. Seventy-three were pursuing a college certificate or diploma, 97 an undergraduate degree, 19 a university certificate or diploma, 26 a graduate degree and 5 students pursued other credentials. Students had a mean age of 27 (SD = 7.82, range = 19–59, median =25). Numbers do not always total the entire sample because some participants did not answer all questions.

Table 1 shows that the most common impairment (41%) noted by students was a learning disability, followed by mobility impairment (23%). Of the 223 students, 100 (45%) reported more than one impairment for a total of 381 impairments. Students were enrolled in a variety of disciplines such as arts, science, nursing, engineering, law, and physiotherapy.

<table>
<thead>
<tr>
<th>Disability/Impairment</th>
<th>Percentage of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning disability</td>
<td>41%</td>
</tr>
<tr>
<td>Mobility impairment / wheelchair user</td>
<td>23%</td>
</tr>
<tr>
<td>ADD/ADHD</td>
<td>21%</td>
</tr>
<tr>
<td>Psychological/psychiatric disability</td>
<td>17%</td>
</tr>
<tr>
<td>Health / medically related impairment</td>
<td>16%</td>
</tr>
<tr>
<td>Deaf / hard of hearing</td>
<td>13%</td>
</tr>
<tr>
<td>Difficulty using hands and/or arms</td>
<td>12%</td>
</tr>
<tr>
<td>Visual impairment: low vision</td>
<td>11%</td>
</tr>
<tr>
<td>Neurological impairment</td>
<td>11%</td>
</tr>
<tr>
<td>Speech or communication impairment</td>
<td>3%</td>
</tr>
<tr>
<td>Totally blind</td>
<td>2%</td>
</tr>
<tr>
<td>Other</td>
<td>1%</td>
</tr>
</tbody>
</table>

*Note.* The 223 students indicated 381 impairments.

Thirty-three campus disability service providers worked at a junior/community college and 23 at a university. One worked at another type of postsecondary institution. They had worked providing services an average of 7 years (range = 1–22 years, SD = 5, median = 5). They indicated that an average of 280 students (SD = 309, range = 4–1100) were registered to receive services from their office. However, there was a very large range and the median of 130 is probably a better measure of central tendency than the mean in this case.

All professors had taught at least one student with a disability during the last three years in a course where they used some form of e-learning. Seventeen worked at a junior/community college and 10 at a university. One failed to specify a school. They taught in a variety of disciplines from arts, science and commerce, to communications and millwright technology. They had taught an average of 25 students with disabilities (SD = 84, range 1–450, median = 6) during the past three years. Given the great range and the variability, the median of 6 students probably represents a better central tendency. They had been teaching in postsecondary education for a mean of 10 years (SD = 7, range = 1–25 years, median =11). Professors had used e-learning for a mean of 5 years (SD = 3.13, range = 1–12 years, median =5).

Seventeen e-learning professionals worked at a junior/community college and 16 at a university. They worked at 32 postsecondary institutions for an average of 18 years (SD = 9, range 2–35, median 20 years), eight of these in the e-learning field (SD = 4, range 1–16, median 8). Their job titles were varied and included Help Desk Intern, Distance Education Specialist, and Director of Educational Media Development.

**Procedure**

**Participant Recruitment**

Four methods were used to recruit potential participants: email discussion lists focusing on Canadian postsecondary education or e-learning, our project partners (National Educational Association of Disabled Students, Canadian
Association of Disability Service Providers in Postsecondary Education, and the Adaptech Research Network), reaching out to previous research participants who expressed willingness to participate in future studies and, in the cases of professors and e-learning professionals, where we had challenges finding participants, we asked our campus disability service provider contacts to recommend potential participants, whom we contacted by phone or email. In all cases, participants who indicated their interest were directed to a website where they read the consent form which provided information about the study, including the draw for four $100 gift certificates to a large online computer store for each group. Clicking on the "I consent" button brought participants to the online questionnaire for their specific group.

**Online Questionnaire**

This was administered in the first half of 2006. The online package, one each for the four participant groups, consisted of demographic questions, closed-ended e-learning related questions required for the larger investigation, open-ended questions of the present study, and a coupon to enter the draw for the gift certificate. Questions were developed in consultation with the project team and project partners. Questions were pilot tested to uncover problems. These were also tested to ensure they were technologically accessible by key informant students with different disabilities. Once complete, we conducted four-week test-retest reliability testing of the closed-ended questions. Questions with poor reliability were omitted from the final questionnaires, which are available from the first author.

**Demographic questions.** We asked all participants about their sex and the name of their school. We also asked: students about their field of studies, the nature of their disabilities, and the adaptive computer technologies they used; professors about their discipline, the numbers of years they taught in postsecondary education, and the number of students with disabilities they had taught during the past 3 years; disability service providers how long they had worked providing services to postsecondary students with disabilities; and e-learning professionals how long they had worked in postsecondary education and their job title. We have used most of these questions in previous studies (Fichten, Barile, & Asuncion, 1999; Fossey et al., 2005).

**Definitions of e-learning and accessibility.** E-learning was defined in the questionnaires as referring to the range of information and communication technologies that professors use when teaching their courses entirely in the classroom, entirely online, or a combination of both. E-learning includes (but is not limited to) the use of the Internet (e.g., course web pages, lectures delivered live online), CD-ROMs, and presentation tools (e.g., PowerPoint). Accessibility was defined as incorporating usability concepts and referring to the ability of students, regardless of their disability, to easily and independently use e-learning. We noted that for some learners this may require adaptive technology (e.g., software that reads what is on the screen).

**E-learning benefits, problems and solutions.** We asked students to write, in free-form text boxes, three benefits of using e-learning. We also asked all participants to respond to the following questions, "Indicate three problems you have encountered using e-learning. How was each resolved? (if not resolved write 'unresolved')" and provided them with three pairs of free-form text boxes.

**Coding problems, solutions, and benefits questions.** A coding manual consisting of 28 categories of Problems, 18 categories of Solutions, and 21 categories of Benefits of using e-learning was developed, based on the open-ended responses that participants provided. Three trained research assistants classified responses. Inter-rater agreements for Problems, Solutions, and Benefits based on five random spot-checks of reliability ranged from 61% to 87%. All discrepancies were resolved by consensus.

**Results**

**Adaptive Technologies**

One hundred thirty-nine students (62%) indicated that they required adaptive technology (i.e., adaptive hardware and/or software) to use e-learning effectively (e.g., software that improves writing quality, screen reader, dictation
There was no significant difference in the number of e-learning problems indicated by students who did and those who did not use adaptive technologies ($M = 2.12$ and $M = 2.22$, respectively, $t(167) = .68$, $p = .50$).

The 139 students who required adaptive technologies to use e-learning reported 368 different technologies, indicating a mean of 2.65 different technologies per student. Table 2 shows that the most commonly used technology was software that helps with writing followed by software that reads what is on the screen. Table 2 also shows that most students felt they could use their adaptive technologies effectively, although there was substantial variability in responses.

<table>
<thead>
<tr>
<th>Benefit Category</th>
<th>Percentage of Students</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of online course notes</td>
<td>41%</td>
<td>87</td>
</tr>
<tr>
<td>Helps learning/academic success</td>
<td>21%</td>
<td>46</td>
</tr>
<tr>
<td>Helps understand course lecture/content</td>
<td>20%</td>
<td>42</td>
</tr>
<tr>
<td>Benefit other than listed</td>
<td>19%</td>
<td>41</td>
</tr>
<tr>
<td>Ability to work at own pace</td>
<td>18%</td>
<td>38</td>
</tr>
<tr>
<td>Ability to work/learn from home</td>
<td>17%</td>
<td>37</td>
</tr>
<tr>
<td>Availability of online course materials/resources other than notes</td>
<td>16%</td>
<td>34</td>
</tr>
<tr>
<td>Helps organization/time management</td>
<td>15%</td>
<td>33</td>
</tr>
<tr>
<td>Convenience communicating with peers/professors</td>
<td>15%</td>
<td>32</td>
</tr>
<tr>
<td>Availability of information anywhere / any time</td>
<td>10%</td>
<td>22</td>
</tr>
<tr>
<td>Feel more independent / confident / less stressed</td>
<td>9%</td>
<td>19</td>
</tr>
<tr>
<td>Saves time</td>
<td>9%</td>
<td>19</td>
</tr>
<tr>
<td>Allows use of adaptive technology/software</td>
<td>8%</td>
<td>18</td>
</tr>
<tr>
<td>Ability to keep up with the rest of the class</td>
<td>6%</td>
<td>13</td>
</tr>
<tr>
<td>Makes classes more interesting and stimulating</td>
<td>5%</td>
<td>10</td>
</tr>
<tr>
<td>Availability of materials in alternate formats</td>
<td>4%</td>
<td>8</td>
</tr>
<tr>
<td>Learn more about technology</td>
<td>4%</td>
<td>8</td>
</tr>
<tr>
<td>Ability to be anonymous / reduces social anxiety</td>
<td>3%</td>
<td>7</td>
</tr>
<tr>
<td>Less materials to transport</td>
<td>3%</td>
<td>7</td>
</tr>
<tr>
<td>Saves money</td>
<td>2%</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3. Percentage of Students with Disabilities Reporting Each Benefit Category

<table>
<thead>
<tr>
<th>Benefit Category</th>
<th>Percentage of Students</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of online course notes</td>
<td>41%</td>
<td>87</td>
</tr>
<tr>
<td>Helps learning/academic success</td>
<td>21%</td>
<td>46</td>
</tr>
<tr>
<td>Helps understand course lecture/content</td>
<td>20%</td>
<td>42</td>
</tr>
<tr>
<td>Benefit other than listed</td>
<td>19%</td>
<td>41</td>
</tr>
<tr>
<td>Ability to work at own pace</td>
<td>18%</td>
<td>38</td>
</tr>
<tr>
<td>Ability to work/learn from home</td>
<td>17%</td>
<td>37</td>
</tr>
<tr>
<td>Availability of online course materials/resources other than notes</td>
<td>16%</td>
<td>34</td>
</tr>
<tr>
<td>Helps organization/time management</td>
<td>15%</td>
<td>33</td>
</tr>
<tr>
<td>Convenience communicating with peers/professors</td>
<td>15%</td>
<td>32</td>
</tr>
<tr>
<td>Availability of information anywhere / any time</td>
<td>10%</td>
<td>22</td>
</tr>
<tr>
<td>Feel more independent / confident / less stressed</td>
<td>9%</td>
<td>19</td>
</tr>
<tr>
<td>Saves time</td>
<td>9%</td>
<td>19</td>
</tr>
<tr>
<td>Allows use of adaptive technology/software</td>
<td>8%</td>
<td>18</td>
</tr>
<tr>
<td>Ability to keep up with the rest of the class</td>
<td>6%</td>
<td>13</td>
</tr>
<tr>
<td>Makes classes more interesting and stimulating</td>
<td>5%</td>
<td>10</td>
</tr>
<tr>
<td>Availability of materials in alternate formats</td>
<td>4%</td>
<td>8</td>
</tr>
<tr>
<td>Learn more about technology</td>
<td>4%</td>
<td>8</td>
</tr>
<tr>
<td>Ability to be anonymous / reduces social anxiety</td>
<td>3%</td>
<td>7</td>
</tr>
<tr>
<td>Less materials to transport</td>
<td>3%</td>
<td>7</td>
</tr>
<tr>
<td>Saves money</td>
<td>2%</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2. Adaptive Computer Technologies Used by Students in Rank Order

<table>
<thead>
<tr>
<th>Adaptive Computer Technology</th>
<th>Percent of Students</th>
<th>Number of Students</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software that helps with writing (e.g., Inspiration, WYNN)</td>
<td>60%</td>
<td>83</td>
<td>5.12</td>
<td>1.36</td>
</tr>
<tr>
<td>Software that reads what is on the screen (e.g., JAWS, ReadPlease)</td>
<td>50%</td>
<td>69</td>
<td>5.31</td>
<td>1.02</td>
</tr>
<tr>
<td>Scanning and optical character recognition (OCR)</td>
<td>35%</td>
<td>49</td>
<td>4.78</td>
<td>1.75</td>
</tr>
<tr>
<td>Voice dictation software</td>
<td>34%</td>
<td>47</td>
<td>4.56</td>
<td>1.56</td>
</tr>
<tr>
<td>Software that magnifies what is on the screen</td>
<td>24%</td>
<td>34</td>
<td>5.03</td>
<td>1.47</td>
</tr>
<tr>
<td>Large screen monitor</td>
<td>24%</td>
<td>33</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Adapted mouse</td>
<td>14%</td>
<td>19</td>
<td>4.25</td>
<td>2.11</td>
</tr>
<tr>
<td>Adapted keyboard</td>
<td>13%</td>
<td>18</td>
<td>4.35</td>
<td>1.93</td>
</tr>
<tr>
<td>Other</td>
<td>7%</td>
<td>11</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Refreshable Braille display</td>
<td>4%</td>
<td>5</td>
<td>4.00</td>
<td>2.74</td>
</tr>
</tbody>
</table>

1 139 students reported using at least one adaptive computer technology; several indicated using more than one.

2 Six point scale, with higher scores indicating being able to use the technology more effectively.
Benefits

The 214 students who responded to this question indicated 578 Benefits. Table 3 shows the percentage of students who indicated each Benefit in rank order. This shows that the most common benefit, noted by 41% of students, was the availability of online course notes.
Problems and Solutions

Two hundred eighty-four participants indicated at least one problem: 184 students, 57 disability service providers, 21 professors, and 22 e-learning professionals. Their responses are shown in Figure 1, which presents the percentage of participants from the four groups who indicated each of the 28 types of problems, in rank order of student responses. It should be noted that student responses are influenced by the disproportionate number of students with learning disabilities. Problems specified by at least 10% of students were, in rank order: inaccessibility of websites/course management systems, technical difficulties, poor use of e-learning by professors, difficulty connecting to websites/course management systems, and students' lack of knowledge of how to use e-learning. Figure 1 also shows that several problem areas were identified by over 10% of the other three groups, although these were noted by less than 10% of students. For example, although 19% of professors indicated problems with PowerPoint presentations in class, less than 5% of students indicated this. Similarly, while 17% of professors indicated problems with time limits on online exams, only 6% of students singled this out as problematic.

![Figure 2. Percentages of Participants Reporting Each Solution](image-url)
Chi Square tests of independence, detailed in Figure 1, show significant findings on five comparisons: technical difficulties (frequently noted by students but not other groups), inaccessibility of course materials (e-learning professionals were most likely to think this is a problem and disability service providers least likely), professors' lack of knowledge about how to make e-learning accessible (students were least likely to think this was problematic), lack of technical support (students and e-learning professionals did not see this as problematic) and inaccessibility of computer facilities (professors were more likely to see this as problematic than the other groups).

Two hundred sixty-eight participants provided at least one Solution: 181 students, 55 disability service providers, 20 professors, and 22 e-learning professionals. Their responses are shown in Figure 2, which presents the percentage of participants from the four groups in rank order of student responses. The most common solutions, noted by at least 10% of students, are: that the problem was not resolved, that the problem was solved using a non e-learning solution, that the professor provided assistance, and that the student obtained needed computer software or hardware. It should be noted that the most common response of all groups was "unresolved," with 67% of students indicating that at least one of their three e-learning problems remained unresolved.

Chi Square tests of independence, detailed in Figure 2, show significant findings on five comparisons: problem unresolved (more frequently noted by students and professors than the other two groups), student obtaining alternate formats (least frequently noted by students), students obtaining adaptive technology (most frequently noted by e-learning professionals), time extensions for online exams (least frequently noted by students and disability service providers), and the school developing accessibility guidelines (noted exclusively by disability service providers). Figure 2 also shows that several solutions were identified by over 10% of one the other three groups, although these were noted by less than 10% of students. For example, although 11% of disability service providers indicated that a campus based disability service provider provided assistance, only 4% of students indicated this. Similarly, although 14% of e-learning professionals indicated that an e-learning professional provided assistance, only 8% of students, 2% of disability service providers and none of the professors cited this as a solution.

![Table 4. Problem-Solution Pairs for Top Five Problems of Each Group](image)

One question of interest concerned the solution for specific problems. Table 4 presents results which show that most of the common problems noted by participants, with the exception of professors, remained unresolved. This was true
for ALL of the students' top five problems, for 4 of disability service providers' and ½ of e-learning professionals' top five items.

When it comes to e-learning problems and solutions the nature of students' disabilities and impairments can have an important impact. Therefore, in Table 5 we present the most common problems and solutions for students with different disabilities. This shows that the most popular solution for students with all types of disabilities is unresolved. For most groups of students, solving e-learning problems by using non e-learning solutions was also popular. In addition to the common problems of inaccessibility of websites and course management systems and technical difficulties, which seem to pose problems for students regardless of the nature of their disability, students with learning disabilities and students with mobility impairments and arm/hand issues also had problems due to their lack of knowledge about how to use e-learning effectively. Students with psychiatric and with health issues noted problems due to poor use of e-learning by professors. Students with hearing impairments, not surprisingly, had problems related to the accessibility of audio and video materials. Students with visual impairments had problems related to the accessibility of course notes and materials, especially those in PDF. When their problem had a solution it was through non e-learning solutions, such as having someone read the materials aloud to them or through alternative formats or using adaptive technologies.

### Table 5. Common Problems and Solutions for Students with Different Disabilities

<table>
<thead>
<tr>
<th>Student's Disability / Impairment</th>
<th>Problems in Rank Order</th>
<th>Solutions in Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning disability</td>
<td>Technical difficulties</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Students' lack of knowledge of how to use eLearning</td>
<td>Non-eLearning solution</td>
</tr>
<tr>
<td></td>
<td>Unresolved</td>
<td>Professor provided assistance</td>
</tr>
<tr>
<td>Mobility impairment / Difficulty using hands and/or arms</td>
<td>Inaccessibility of websites / course management systems</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Technical difficulties</td>
<td>Non-eLearning solution</td>
</tr>
<tr>
<td></td>
<td>Students' lack of knowledge of how to use eLearning</td>
<td>Student obtained / used technology / software</td>
</tr>
<tr>
<td>ADD/ADHD</td>
<td>Technical difficulties</td>
<td>Unresolved</td>
</tr>
<tr>
<td>Psychological / psychiatric disability</td>
<td>Poor use of eLearning by professors</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Difficulty connecting to websites / course management systems</td>
<td>Obtained alternate formats</td>
</tr>
<tr>
<td>Health / medically related impairment</td>
<td>Technical difficulties</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Poor use of eLearning by professors</td>
<td>Non-eLearning solution</td>
</tr>
<tr>
<td></td>
<td>Inaccessibility of websites / course management systems</td>
<td></td>
</tr>
<tr>
<td>Deaf / hard of hearing</td>
<td>Inaccessibility of audio/video material</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Unresolved</td>
<td>Professor provided assistance</td>
</tr>
<tr>
<td>Visual impairment: low vision</td>
<td>Inaccessibility of course notes/materials</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Inaccessibility of websites / course management systems</td>
<td>Non-eLearning solution</td>
</tr>
<tr>
<td></td>
<td>Inaccessibility of course notes/materials in PDF</td>
<td>Obtained alternate formats</td>
</tr>
<tr>
<td>Totally blind</td>
<td>Inaccessibility of websites / course management systems</td>
<td>Unresolved</td>
</tr>
<tr>
<td></td>
<td>Inaccessibility of course notes/materials in PDF</td>
<td></td>
</tr>
</tbody>
</table>

Discussion

**Limitations of the Present Investigation**

There were several limitations. While the number of participants is large and diverse, the samples are neither random nor, we believe, fully representative of the populations studied. Given self-selection biases, and the fact that recruitment was carried out primarily online, we expect that the proportion of students who read online discussion lists and have experience using e-learning are over represented. Similarly, disability service providers passionate about and/or more heavily engaged in e-learning and accessibility are likely to be over sampled. And we deliberately sought out specific professors and e-learning professionals who had had experience with e-learning and students with disabilities. Even more troubling, we are unable to calculate a final "return rate" because of the manner in which we recruited participants. The use of email discussion lists as the main form of recruitment poses challenges in this regard.
Yet, available indices suggest that our samples have characteristics which resemble the realities of Canadian postsecondary education. The age range of students is normative for studies of students with disabilities (e.g., Henderson, 1999; Horn & Berktold, 1999; Killean & Hubka, 1999; Middleton, 2003). The sample contains more female than male students and disability service providers; this too, is characteristic of these populations in postsecondary institutions (e.g., Fossey et al., 2005; Sharpe et al., 2005; Michaels, Prezant, Morabito, Jackson, 2002; Statistics Canada, 2007).

**Findings Unique to Students**

**Types of Disabilities/Impairments**

The results indicate that, consistent with the findings of others (e.g., Stodden, 2006), the most common disability indicated by students was a learning disability, often coupled with attention deficit disorder/attention deficit hyperactivity disorder. Our sample also reflects the trend toward the "new clientele" of students with disabilities in postsecondary education (Fichten et al., 2006; Fiset, 2006), with many students reporting psychological/psychiatric disabilities and health/medically related impairments. Almost half of the students reported more than one disability, a finding consistent with the substantial number of students who have multiple impairments reported in the literature (e.g., Fichten et al., 2006).

**Adaptive Technology Use**

Over half of the sample indicated that they needed some form of adaptive technology, such as software that improves writing quality, screen reader, or dictation software to access e-learning effectively. Two-thirds of the students who used adaptive technology indicated using more than one type, suggesting the need to ensure that these can work with each other. Given the nature of students' impairments it is not surprising that the most common software that students indicated is software which helps with writing quality (e.g., WYNN, TextHelp) followed by software that reads what is on the screen (e.g., ReadPlease, Jaws), scanning and optical character recognition software (OCR - both specialized products for students with visual impairments, such as Open Book, and general use products such as OmniPage), and voice dictation software (e.g., Dragon Naturally Speaking).

**Benefits of e-Learning**

Consistent with the results of others (e.g., Goodman, Tiene, & Luft, 2002), students were enthusiastic about the benefits of e-learning. The most popular response, noted by over 40% of students, was the availability of online course notes. Students also noted the benefits of information anywhere and any time and the availability of online course materials other than course notes. Students also noted that e-learning allows them to work at their own pace, to study from home, and to easily communicate with peers and professors.

**Problems and Solutions**

**Commonalities Among Problems**

All groups complained of the inaccessibility of websites and course/learning management systems (CMS). This was the most common problem reported by all four groups of participants. Problems reported with these is hardly surprising given that these are the most popular means of delivering e-learning in postsecondary education (Malik, Asuncion, & Fichten, 2005). Of course, this type of problem can reflect both accessibility and usability issues with CMS systems, such as WebCT, as well as problems with course websites developed by individual professors, departments and schools. In addition, problems experienced can be due both to software issues as well as to content put into a CMS and websites. Content is often designed by professors, departments, and schools, as well as by e-vendors such as the producers of course textbooks. So difficulties in this realm could be due, for example, to accessibility issues directly related to the CMS (e.g., a chat interface which cannot be accessed using dictation software); to usability issues such as the 10 frames inserted by a course text book publisher into a WebCT shell; and
to inaccessible content, such as image based PDFs which cannot be read by screen reading software and video clips without captioning (subtitles). Other issues noted by participants included difficulties using some websites by screen reading technologies, confusing structures for web pages for students with learning disabilities, problems with fixed font size of materials on websites, and problems for students with visual impairments to access online maps and images.

Another problem area identified by all groups relates to accessibility of audio and video materials. Online and textbook-related video clips as well as "taped" lectures are becoming commonplace. Although most free computer based video players (e.g., Windows Media Player) are capable of displaying closed captions (similar to subtitles in movies), most digital videos needed by students are not captioned, making them inaccessible to students with hearing impairments. The absence of a description about what is going on in the video causes accessibility problems for many students with visual impairments. The same is true for online lectures and other multimedia materials which are generally not captioned, described, or accompanied by written materials such as a text file. This type of problem also arises in video and audio chats.

All groups also found inflexible time limits to complete activities built into online exam and assignment CMS products problematic. In many schools, students with a variety of disabilities are entitled to additional time to complete the tests and assignments (e.g., Harding, Blaine, Whelley, & Chang, 2006; Sharpe, et al., 2005). But faculty can usually specify only one time duration for all students in most online evaluation systems, making the allocation of additional time impossible.

PowerPoint and data projection during lectures can also pose accessibility problems for students, mainly for those with visual and hearing impairments. Students with visual impairments, of course, experience difficulty seeing the presentation while those with hearing impairments have difficulty following the onscreen presentation and taking notes at the same time. These students also have problems viewing both the presentation as well as the professor's lips (or the sign language interpreter) at the same time, especially with the lights off.

Another area identified by all groups concerns problems with course notes and materials in PDF. The problem with PDF is that its accessibility is dependent on how the PDF was made. Faculty often scan old, heavily annotated documents and articles and distribute these to students in PDF format. Scanning produces a PDF document consisting of "images" of the scanned pages — these cannot be edited and sections cannot be copied and pasted. Students who use screen reading technologies cannot access these documents since there are no words to read. Such PDFs need to be rendered into electronic text by optical character recognition (OCR) software. But this works very poorly if the original paper version had been heavily annotated or underlined or if it had already been photocopied several times. Similarly, unless specifically marked-up to be accessible, documents with multiple columns and those with tables and figures, even when made into PDF from clean originals, can create difficulties because of the way in which screen readers handle this kind of text.

All groups also identified lack of needed adaptive technologies as problematic. This can be missing in the school's specialized as well as general use computer labs and can leave students with disabilities merely sitting in class while their peers are engaged with the e-learning materials provided by the professor. Participants noted problems such as inadequate numbers of adapted keyboards, insufficient site licenses for adaptive computer technologies, and difficulties caused by the inability to install adaptive software on locked work stations.

Problems Unique to Each Group

In many cases, some but not all groups identified a specific item as problematic. Best seen in Figure 1, these generally reflect the respondent's perspective (e.g., end-user vs. campus professional). Disability service providers, whose role sees them responsible for accommodating students with disabilities and acting as an advocate for them on campus, are most likely to deal with problems that students, sometimes in collaboration with their professors, cannot resolve. Problems they face are likely to be different from those confronting students or the other groups. For example, they may find problems experienced by students who are blind as especially problematic, even though there are few students with this disability on campus. Professors generally experience problems in the context of their own courses and students. So the types of problems and solutions they indicated reflect the types of e-learning they use and the disabilities of students in their classes. Similarly, e-learning professionals are more likely to
experience systemic problems and technical issues that the other groups are not familiar with. They also have limited
direct contact with students in general, and with students with disabilities in particular.

Students described technical difficulties using e-learning and problems connecting to websites and course
management systems. They also had problems downloading and opening files, they had difficulties with web pages
that would not load, and video clips that took long times to download. They also noted poor use of e-learning by
professors. Commonly reported issues were: the professor does not get the material onto the website on time; course
notes online are incomplete; in class PowerPoint presentations are delivered too quickly. Students also indicated that
they themselves also lack knowledge about how to use e-learning (e.g., not knowing how to log in or use the CMS,
finding web searches for learning materials complicated, needing to learn to use online learning technologies). Many
of these problems are probably shared by nondisabled students and research comparing e-learning problems and
solutions of students with and without disabilities is needed for comparison.

Disability service providers, who are generally responsible for arranging accommodations for students with
disabilities, also wrote about poor use of e-learning by professors. They also identified problems with poor
accessibility of course notes and materials in formats other than PDF. For example, interactive mathematics graphs,
inaccessible maps and images, and problematic supplementary materials provided by e-vendors such as CD based
crossword puzzles and online flash cards and practice tests.

Professors noted problems that are likely to be brought to their attention by their students. These were all raised by
members of the other groups.

E-learning professionals, perhaps because they are the ones to whom all others turn when other avenues have been
exhausted, failed to mention many of the problems raised by students and disability service providers. They did,
however, note difficulties with inaccessible course notes and materials (36% noted this as a problem).

**Solutions**

The results show that the most common response, for all four groups, was "unresolved" with 67% of students, 53% of
disability service providers, 36% of e-learning professionals and 35% of professors indicating that at least one of
their three e-learning problems remained unresolved. Solving an e-learning problem with a non e-learning solution
(e.g., taking course materials home to have the student's husband read the materials aloud, writing an exam at a
different time from the rest of the class) ranked next in popularity for all groups except e-learning professionals. The
student obtaining alternate formats was a popular response for most groups, except for students, who were least
likely to note this as a solution to their problem.

Again, certain solutions were noted by some but not all groups of respondents. For example, the professor providing
assistance was mentioned relatively frequently by both students and professors. E-learning professionals indicated
that an e-learning staff member provided assistance (e.g., fixing logins, altering time frameworks for exams). Disabilty
service providers indicated that a campus disability service provider provided assistance. Because of their
vantage point, professors, disability service providers, and e-learning professionals are more likely to be aware of
their own contribution to solving problems than the contributions of others (Fichten, 1984).

Obtaining or using adaptive technology was seen as solving the problem by a large proportion of e-learning
professionals, but by virtually none of the other groups. Examples include: use of video equipment that enlarges the
instructor's face to help a student lip read, converting an online exam to an alternate format that was accessible to the
student, and scanning PDFs and distributing the CDs to students. The problem being resolved by providing time
extensions for online exams was a popular option mainly for professors.

**Problem-Solution Pairs**

We wanted to find out whether certain problems were more easily resolved than others. To do this we inspected
solutions to problems noted frequently by each participant group. As expected, most problems remained unresolved.
For problems identified by students where a solution was found the results indicate that inaccessibility of websites and
professors' poor use of e-learning were resolved by professors and that technical difficulties and problems
connecting to websites and course/learning management systems were resolved by e-learning specialists. Disability
service providers’, professors’, and e-learning professionals' problems related to inaccessibility of course notes and PowerPoint and data projection were resolved by obtaining alternate formats.

Problems and Solutions of Students with Different Disabilities/Impairments

Because the nature of students' impairments was also likely to influence problems and solutions we also investigated how problems noted by students with different impairments were resolved. The results show that students with learning disabilities indicated experiencing technical difficulties and problems caused by their own lack of knowledge about how to use e-learning. This was also true for students with mobility impairments and difficulty using their hands and/or arms, but these students also noted problems related to inaccessible course/learning management systems. Students with attention deficit and/or attention deficit hyperactivity disorder experienced technical difficulties most often as did students with health and medically related impairments. But this latter group of students also noted problems with poor use of e-learning by professors and inaccessibility of websites and CMS. Students with psychological/psychiatric disabilities also noted poor use of e-learning by professors as well as difficulty connecting to websites and CMS. It was not surprising to find that students with hearing impairments experienced problems with inaccessibility of digital audio and video materials and that students with visual impairments noted problems with inaccessibility of course notes/materials, especially those in PDF as well as problems related to the accessibility of websites and CMS.

Recommendations

Training

One means of addressing problems involving inaccessibility of websites and course management systems, of e-learning broadly, and of specific materials, such as course notes and audio and video clips is through training of professors. Many colleges and universities already offer training on how to integrate e-learning in teaching and on how to use specific e-learning tools. Developing a module, as a start, on how to make e-learning accessible, and integrating this into existing training, would, at a minimum, begin sensitizing faculty and staff on the issues. Other, more targeted sessions can be considered on specific topics, such as how to make a website or PDF file accessible, based on needs. There are numerous online resources to act as a starting-point (e.g., EASI <http://easi.cc>, DO-IT <www.washington.edu/doit>, and WebAIM <www.webaim.org>).

Similarly, to address students' concerns about their lack of knowledge on how to use e-learning, any existing training opportunities available to students at large on e-learning use should be promoted to students with disabilities as well. It goes without saying that such training would need to be accessible to these students.

Adopt e-Learning Accessibility Guidelines

Colleges and universities should consider developing and adopting e-learning accessibility guidelines that address both in-house development of e-learning as well as purchases of e-learning products and technology. Wording of these should be in language that is readily understood by faculty (Gabrielli, Mirabella, Kimani, & Catarci, 2006). Like training, having such guidelines in place would help resolve problems with inaccessible websites and other e-learning tools and materials, and would inform those making purchasing decisions about the need to select the most accessible product. Of course, it goes without saying that the strength of such guidelines would be based on the commitment demonstrated by those who lead and/or champion e-learning on campus.

Proactively Engage On-campus Accessibility Experts

Whether it is testing a new e-learning tool with end-users or running a committee whose focus is e-learning, ensure that accessibility is represented. In the case of end-user testing, actively seek out and invite students with different disabilities to participate in such activities. Who better to identify possible accessibility issues than the users themselves? Such users can provide first-hand insight into potential accessibility problems, particularly if they require the use of adaptive hardware or software to interact with the e-learning. In the case of committees, engage,
where possible, someone who is accountable for supporting students with disabilities so that accessibility interests are on the table. At a minimum, make sure that accessibility is a standing agenda item, so that any issues can be raised on a regular basis. Both of these activities not only increase the visibility and awareness of accessibility in the e-learning context, but they also help bring to light problems such as those raised by study participants.

The fact that “unresolved” was the most frequently noted solution to addressing e-learning accessibility problems by all four groups of participants cannot go unmentioned. While not within the scope of this study, it is critical to understand the implications of this finding given the prevalence of e-learning use in colleges and universities today. How is this impacting on the ability of students with disabilities to succeed academically and to compete and participate on a level playing-field with their non-disabled peers? Why do problems remain unresolved? Is it the reluctance of institution’s to take steps to make accessibility a requirement for professors? Is it poor accessibility of the products produced by e-vendors, such as software developers and book publishers? Or is there a knowledge gap that needs to be closed? These are fundamental questions that postsecondary institutions have a shared responsibility to answer.

Conclusions

During the last decade there has been tremendous development and interest in e-learning on campus. While our research shows the many benefits of e-learning, such as the availability of online course notes, there are also problems. Chief among these are problems related to inaccessibility of websites and course management systems. Our study also illustrates that problem experiences differ among the four groups of stakeholders in our study. This is most likely attributable to the differing roles and levels of interaction with accessibility of e-learning that each group has. That being said, the most commonly cited response to solving e-learning problems by all four groups was “unresolved.” This is a cause for concern given the prevalence of e-learning use on campus. It is the charge to all postsecondary stakeholders to ensure that e-learning technologies continue to benefit rather than hamper students with all types of disabilities and that accessibility gains are maintained and built upon.

Acknowledgements

Funding for this research was provided by the Social Sciences and Humanities Research Council of Canada (SSHRC), the Dis-IT Research Alliance, and the Canadian Council on Learning. The research was carried out with the collaboration of our partners: the Adaptech Research Network, NEADS, AQEIPS, CADSPPE, and AQICEBS.

References


Can Online Peer Assessment be Trusted?

L’hadi Bouzidi¹ and Alain Jaillet²

¹Department of Mathematics, Exact Science Faculty, A. Mira University, Bejaia, Algeria // lhadi_bouzidi@yahoo.fr
²ULP Multimedia, Strasbourg, France // alain.jaillet@ulpmn.u-strasbg.fr

ABSTRACT
The excessive workload generated by the assessment of exam papers in large classes and the need to give feedback in time often constitute a rather heavy burden for teachers. The online peer assessment can contribute to reduce this workload and, possibly, to improve learning quality by assigning the assessment task to students. However, this raises the question of validity. In order to study this question, we carried out an experiment of online peer assessment in which 242 students, enrolled in 3 different courses, took part. The results show that peer assessment is equivalent to the assessment carried out by the professor in the case of exams requesting simple calculations, some mathematical reasoning, short algorithms, and short texts referring to the exact science field (computer science and electrical engineering).

Keywords
Peer Assessment, Self-Assessment, Peer Assessment Validity, Item Analysis

Introduction
Learning assessment often constitutes a problematic task for the teacher. Its repetitive nature makes the teacher's job fastidious and the grading of hundreds of exam papers is far from exciting. In addition, the educational value of the assessment methods is often very low. The post-exam phases are difficult to put into practice with a large number of students, particularly, the feedback phase. Moreover, students are not really fond of these phases. In order to improve this situation, several devices can be used depending on the aspects that need to be changed. In this study, our primary concern is not to improve the reliability or the quality of assessment (De Ketele & Gerard, 2005), but rather to reduce the teacher's workload, without affecting the quality of assessment. In order to attain this, we suggest the use of computer-assisted assessments (Davies, 2001; Stephens & Bull, 1998) and the application of online peer assessment (Topping, 1998). Our proposal could be acceptable for the formative assessment of the learning process (Perrenoud, 2001) as this assessment method does not have many direct consequences on decisions concerning access to the following year or to certification. However, what about summative evaluation? Can grading carried out by students be trusted? What are the conditions of acceptability of our proposal? Several attempts (Topping, 1998) have been made to involve more students in this kind of approach, but it seems that the use of technologies can facilitate the implementation of these solutions. What remains to be established is whether this study can show that the assessment, thus produced, will be acceptable. Our research focuses on this issue of acceptability by measuring precisely the difference between assessments carried out by peers and those carried out by the teacher.

Online peer assessment
The use of online peer assessment is recent. It allows the students to assess the work of their peers and even assess their own work. Unlike the self-assessment techniques, which are generally limited to basic cognitive levels (Bloom, 1956; Anderson & Krathwohl, 2001), peer assessment enables to develop learning at high cognitive levels. This is a method that involves the students in the revision, assessment, and feedback process of work online. Doiron (2003) indicates that certain authors criticize the use of information and communication technology for peer assessment, arguing that it is not as rigorous as traditional types of assessment, that it requires too much student effort by putting too much pressure on them, that it is not reliable, and that it is not necessarily fair. In response, other authors such as Bostock (2000) are convinced that the (formative or summative) assessment of other students' work by the students themselves has several advantages for the learning process, both for the assessor and the assessee. Bostock points out that peer assessment encourages the students to be independent and develops skills in high cognitive areas. He acknowledges certain weaknesses of this type of assessment, in particular the over-estimating of friends' work, but he explains that this can be avoided by setting up a system which would guarantee anonymity, multiple assessments, a great number of assessors, and moderation by the teacher. Moreover, Bostock, specifies that Internet and information and communication technologies enable an easier management of a greater number of students.
Several authors (Rada & Michailidis & Wang, 1994; Fisher, 1999) explain that peer assessment may be proposed as an alternative solution to reducing teacher workload. Thus, Cho et al. (2006) explain that, despite the progress made during the last two decades in student writing skills, the courses rarely include real comprehensive writing tasks. According to them, this is due to the teachers’ workload: assessing writing skills requires too much time and effort (Rada et al., 1994). They suggest resorting to peers in order to assess the students' work rather than systematically resorting to evaluation by teachers (Rada et al., 1994).

As regards the reliability of peer assessment, few studies have been carried out (Haaga, 1993; Falchikov, 1986; Falchikov & Goldfinch, 2000; Mowl & Pain, 1995; Stefani, 1994; Cheng & Warren, 1999). The majority of these studies have analyzed the performance in only one class (from 45 to 63 students) (Cho et al., 2006). It seems that only one study examined several courses with a significant number of participants (708) (Cho et al., 2006). Various measurements were used: comparison of means, Pearson’s product moment correlation, concordance percentage, intra-class correlation, etc. The results are contradictory. There are good reasons for both confirming and contesting the reliability of marks awarded by peers. Hereafter, we will reveal the results of some of these studies.

Sadler & Good (2006) carried out a peer assessment experiment on 4 secondary school classes. By comparing the marks awarded by teachers with those awarded by pupils and with those obtained by self-assessment, they claim to have obtained a very high correlations ($r = 0.91$ to $0.94$). They also noticed that pupils slightly underestimate work of their comrades, whereas they overestimate their own. Regarding the impact of peer assessment on the learning process, they claim that self-assessment reinforces pupils’ learning whereas peer assessment does not. Lastly, they explain that peer assessment considerably reduces the teachers' workload.

Zevenbergen (2001) believes that peer assessment is an efficient learning tool which can help mathematics teachers, but that should not be seen as an alternative to teacher’s assessment of pupils’ work.

Cho et al. (2006) carried out a peer assessment experiment on 16 classes (708 students) studying the science of education. The results reveal that peer assessment, based on at least 4 assessors per paper, is extremely reliable and is as valid as the teacher’s assessment.

In summary, the literature review reveals that there are good theoretical reasons both for and against the reliability and validity of peer assessment (Cho et al., 2006). The previous empirical work is not large. Although the literature concludes that peer assessment appears valid (Falchikov & Goldfinch, 2000; Topping, 1998), this validity still needs to be addressed in a larger scale study using a common metric, across many courses and levels of students.

**Reliability and validity issues**

In practice, the assessment can pose validity and reliability problems. This may concern the tests, the assessment tools and the assessors (Brennan, 2001). The validity of a test is regarded as acceptable if the test comprises only items enabling the assessment of competences corresponding to training objectives. The reliability of an assessment tool is acceptable if it enables the accurate observation of the learning competence and makes it possible to give an opinion on this competence. An assessor is regarded as reliable if he or she has a good knowledge of the subject concerned by the assessment and if he or she also shows a good aptitude to making responsible judgments.

Validity and reliability remain debatable issues. Cho et al. (2006) note that validity is sometimes misreported in literature as reliability. This leads us to clarify these two concepts based on the definitions provided by Cho et al. (2006) and Falchikov et al. (2000). These authors regard the reliability of peer assessment as a variable that can be measured by the similarity between the marks given by peers. They consider the validity of peer assessment as a variable that can be measured by the similarity between the marks attributed by peers and by teachers.

**Assumptions and objectives**

We are considering tests and assessment tools that should guarantee validity and reliability. On the one hand, teachers are deemed pertinent assessors and are therefore considered reliable. On the other hand, we know that students are neither reliable nor valid assessors. Consequently, we wanted to find out whether several students taken together could be deemed as a valid “collective” assessor. In order to verify this, we compared the marks awarded by
the teacher and those awarded by the groups of students (the peers). If, on a significant sample and on the basis of a 95% confidence limit, we obtained that the marks awarded by the peers were the same as those awarded by the teacher, we would be able to confirm that the students' assessments could be trusted (i.e. that peer assessment is valid). This study is based on tests with closed or half-open questions (short calculations, mathematical reasoning, writing short algorithms, and drafting short texts).

Methodology

The validity of peer assessment can be verified by comparing peer assessment with the assessment carried out by the teacher (Cho et al., 2006). This assumes that:

- The teacher is a reliable assessor,
- The test items are set up correctly,
- The test constitutes a reliable assessment tool (internal consistency of items),
- The observation and marking tools (computer-assisted assessment tools in our case) are reliable,
- And the marking instructions and scale schemes are clear and coherent.

This has led us to use a method which requires that five hypotheses must be confirmed before checking the validity of peer assessment (see figure 1).

![Figure 1. Validation process of peer assessment](image)

As educational experts, teachers are considered reliable assessors. Proving that peer assessment is equivalent to teacher assessment could validate peer assessment. The validity of the tests can be checked through a comparative analysis of the papers of the same test by several experts (at least three) in the subject area (De Ketele et al., 2005). An item analysis (Wright & Panchapakesan, 1969) can also be used for validating the test. The reliability of the test can be checked by analyzing the internal consistency of its various items (Cronbach, 1951). In the case of computer-assisted assessments, it is more difficult to verify the reliability of the observation and marking tools. This difficulty is more important in the case of peer assessment. Human, technical and technological factors often contribute positively or negatively to this reliability. The marking instructions and scale schemes define the rules for awarding marks. Verifying their validity can be carried out by checking the coherence of the scale schemes (independent questions, accuracy) (Dubus, 2006) and by analyzing the correspondence between the question formulation and the marking instructions.

Constructing quality of test items: This was performed by an item analysis (Nunnally, 1967) although we did not carry out a complete quality analysis of the tests. We used this technique to identify some abnormalities in the test items. The level of these abnormalities enabled us to determine whether the test could be used to check the validity of peer assessment or not. We considered two metrics widely used in docimology: the difficulty index and the discrimination index (Girard, 2001). The comparison of the values of these statistical indexes resulting from marks awarded by peers with those resulting from marks awarded by the teacher could only explain certain abnormalities that will be taken into consideration in our decision regarding the validity of peer assessment.

The difficulty index “p” is the percentage of failure to answer the question. It is equal to the number of students having replied incorrectly, divided by the total number of students. When p is close to 0: this shows an easy question. When it is close to 1: this shows a hard question.
The discrimination index “r” (which varies from -1 to +1) represents the correlation (Pearson) between the answers to the question and the total for the other questions. Discrimination is considered (Girard, 2001) to begin at \( r = 0.20 \). The items of average difficulty (\( p \) from 0.40 to 0.60) maximize discrimination. A discrimination coefficient equal to zero indicates that there is no discrimination. When it is negative, this shows incoherence: the best students fail the question, the weakest ones answer correctly. The Office of Educational Assessment of the University of Washington (Scorepak, 2005) classifies the discrimination coefficient as follows: \( r > 0.30 \) indicates that the question is good; \( r \) between 0.10 and 0.30 indicates that the question is acceptable; \( r < 0.10 \) indicates that the question is bad.

The 4 kinds of item abnormalities that have been used to check the test validity are exposed in Table 1, presented hereafter.

### Table 1. Various situations of abnormal items

<table>
<thead>
<tr>
<th>“p” (difficulty index)</th>
<th>“r” (discrimination index)</th>
<th>Abnormality</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between 0.40 and 0.60</td>
<td>Lower than 0.10</td>
<td>A1: Average difficulty and no discrimination</td>
<td>Generally speaking, questions of average difficulty maximize the discrimination of the students</td>
</tr>
<tr>
<td>.( p )</td>
<td>Between -0.10 and 0.10</td>
<td>A2: High negative discrimination</td>
<td>Negative discrimination means that the better students fail the question and the weaker ones answer correctly.</td>
</tr>
<tr>
<td>Higher than 0.80</td>
<td>Higher than 0.40</td>
<td>A3: Difficult question and good discrimination</td>
<td>Questions that are too easy or too hard should not discriminate the students</td>
</tr>
<tr>
<td>Lower than 0.20</td>
<td>Higher than 0.40</td>
<td>A4: Easy question and good discrimination</td>
<td></td>
</tr>
</tbody>
</table>

**Reliability of the test (Internal consistency of the items):** This is the coefficient \( \alpha \) of Cronbach (Cronbach, 1951; Bouchet, Guillemin, Hoang, Cornette & Briançon, 1996), which is the metric usually used. It is a statistical index that varies between 0 and 1 and enables the assessment of an assessment instrument’s homogeneity (the internal consistency or coherence), made up of a series of items that should all contribute to understanding the level of knowledge or skill on a given theme. This index conveys a degree of homogeneity that increases when its value approaches 1. There is no known statistical distribution that makes it possible to conclude whether Cronbach’s alpha is acceptable or not (McKell, 2000). The empirical limits resulting from the psychometrics used as a reference: In an exploratory study, Cronbach’s alpha is considered as acceptable if it is between 0.6 and 0.8 (Zimmerman, Beverly, Sudweeks, Richard, Shelley, Monte et al., 1990; Evrard, Desmet, Dussaix, Lilien, Pras & Roux, 2003; Nunnally, 1967). The Office of Educational Assessment of the University of Washington (Scorepak, 2005) classifies the reliability coefficient as follows:

- \( \alpha > 0.90 \): Excellent reliability; at the level of the best standardized tests.
- \( 0.80 < \alpha < 0.90 \): Very good for a classroom test.
- \( 0.70 < \alpha < 0.80 \): Good for a classroom test. There are probably only a few items that would require improvements.
- \( 0.60 < \alpha < 0.70 \): Somewhat low. The test requires to be supplemented by other measurements (other tests) to determine the grades. There are probably some items that would require improvements.
- \( 0.50 < \alpha < 0.60 \): Suggests a need for revising the test, unless it is quite short (ten or less items). The test definitely needs to be supplemented by other measures (e.g., more tests) for grading.
- \( \alpha < 0.50 \): Questionable reliability. This test should not count as the main contribution to the course grade and needs to be revised.

**The reliability of the observation and marking tools:** In our experiment, the “workshop” module of the Moodle platform was used. This module is stable and presents a user-friendly environment. However, the marking tool (scale scheme) had a weakness. For the scale scheme that we had defined, the marking tools provided by this workshop were in the shape of option buttons varying from “excellent” to “very poor”. Each button corresponded to a mark that the assessor could choose. This method can lead to errors.
Marking instructions and marking schemes: The marking schemes can comprise additive boxes and subtractive boxes (Dubus, 2006) that may coexist within the same total. The condition so that the marking scheme remains coherent is that all the boxes remain independent from each other, i.e. a subtraction box that has used up its credit cannot start nibbling the points scored in another box. Thus, a true marking scheme does not allow any compensation between sections. The idea of goal analysis, which underlies the construction of marking scales, contradicts the idea of compensation, because a weakness in skill A cannot be compensated by a strength in skill B.

The analysis that we carried out, in this research, does not aim to improve or change the marking scheme, but rather to check whether or not they have had an impact on the quality of the assessment (objectivity). We define the quality of a marking scheme with two characteristics: first, its accuracy and, second, the complete independence of its boxes. The accuracy of the marking scheme refers to its ability to help giving opinions on elementary skills. We think that, in order to guarantee the best objectivity, it is important to supplement the marking scheme by marking instructions. The validity of the marking instructions is defined by the clarity and coherence of the wording used for the items (questions). Their analysis may explain abnormal differences between assessors.

In this study, we will resort to checking the quality of the marking instructions only if the items considered contain abnormal differences between assessors.

Validity of peer assessment: It can be defined by the measurement of the agreement between two judges: the teacher and the peers. For qualitative variables, we can use the Kappa nonparametric test of Cohen (Cohen, 1960). For ordinal variables, we can use the Spearman Rho Coefficient (Kruskal, 1958). For quantitative variables (as in this research), we can use Pearson’s correlation coefficient (Cho et al., 2006; Haaga, 1993) or the t-test (Mowl & Pain, 1995; Cheng and Warren, 1999; Stefani, 1994).

Experiment

We chose to implement a digitalization process to put the exam papers carried out in class online. This choice guaranteed an environment faithful to the exams actually carried out in class. The process that we implemented can be described according to the following stages (see figure 2):
1. Organize an exam in class.
2. Collect and then digitalize the exam papers.
3. Process the digitalized papers to ensure that the students remain anonymous.
4. Place the digitalized papers on the virtual campus for each student.
5. Prepare the activity on the virtual campus by indicating the marking scheme, assessment criteria and various parameters for assessing the marks and the distribution of the papers to the students.
6. Open the assessments (for the teacher and for the peers).
7. Hand-over the marks and the feedback to the students and open the discussion.

![Figure 2. Online peer-assessment of classroom exams](image)
In order to carry out our experiment, we based our research on five initial conditions:
- **Subject area**: exact sciences in order to restrict the scope of interpretation.
- **Exam**: performed in a supervised class in order to avoid any risk of invalidation.
- **Marking scheme**: clearly defined by the teacher.
- **Test**: including more than 9 questions.
- **Average timing of the test**: 2 hours.

The clear definition of the marking scales and the marking instructions were a condition on which we insisted. Moreover, we have recommended providing the exams solutions to the assessors.

In order to ensure that the students take this experiment seriously, we informed them that the quality of their assessment would be marked and that the assessment would be anonymous. Thus, they were invited not only to assess, but to correctly assess anonymous exam papers.

**Data collection and treatment**

There were 242 students (see table 2) of the second and the third year in engineering who took part in the experiment. Three exams were organized. They concerned three courses: computer architecture courses taught during the years 2006 and 2007 (CA-2006 and CA-2007), and the general electrical engineering course taught during the year 2007 (GEE-2007).

<table>
<thead>
<tr>
<th>Exams</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>2006</td>
</tr>
<tr>
<td></td>
<td>2007</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total:</td>
</tr>
</tbody>
</table>

The CA-2006 exam is composed of 17 items involving calculations (7 items) and drafting of short texts (one to five words: 5 items, one to five lines: 5 items). Six of these items enable the assessment of the basic cognitive levels: learning, understanding, and applying (Bloom, 1956; Anderson, 2001). The eleven other items enable the assessment of an additional cognitive level: analyzing.

The CA-2007 exam is composed of 13 items involving calculations (6 items), writing of algorithms (1 item) and drafting of short texts (one to five words: 2 items, one to five lines: 4 items). Three items enable the assessment of the low cognitive level: learning. Eight items enable the assessment of the basic cognitive levels. The two remaining items enable the assessment of an additional level: analyzing.

The GEE-2007 exam is composed of 9 items involving calculations and mathematical reasoning (establishment of a succession of mathematical transformations that lead to an expected formula). All the items enable the assessment of the four first cognitive levels (learning, understanding, applying, and analyzing).

All of the exams are marked out of 20 points. Their items are marked out of 0.5, 1, 1.5, 2, 2.5 and 3 points according to the marking scales. The smallest marks that can be awarded are 0.25 and 0.5 according to the exam.

For each test and for each paper, we gathered the following marks (see figure 3): one self-assessment mark, four marks awarded by peers, and one mark awarded by the teacher; this applies for each item. Once this data was gathered, we added up the marks linked to the items (questions) in order to calculate the marks for all of the assessors:
- “TM”: Marks awarded by the teacher.
- “SM”: Marks produced by self-assessment.
- “PM”: Marks awarded by peers disregarding the self-assessment.
- “SPM”: Marks awarded by peers, incorporating the self-assessment.
For each test and for each type of assessment (self-assessment, peer assessment, peer assessment incorporating self-assessment, and assessment by the teacher), we have performed an item analysis. In particular, we have calculated, for each type of assessment and for each test, the following indicators: averages and standard deviations of the marks and Cronbach’s coefficient $\alpha$. We have also calculated the difficulty index and the discrimination index for each item.

**Analysis and interpretation**

In relation to CA-2006 exam, the item analysis produced the results shown in Table 3. We notice an “A3 abnormality” (see Table 1) in Q16 and Q17 questions. However, for these questions, we observe a good agreement between the teacher's assessments and the peers’ assessments. We also note that it is the peer assessment incorporating the self-assessment that is the closest to the assessment by the teacher. The students, with a positive average difference between their marks and those of the teacher (0.12), very slightly under-mark the papers of their peers. On the other hand, the authors of the papers, with a negative average difference between their marks and those of the teacher (-0.83), over-mark their papers.

The difficulty of the other questions, according to the teacher, is generally equal to that found by the peers (correlation equal to 0.97). However, there is an incoherence in the Q14 question that shows a difference of 26% between the teacher’s severity and that of the peers (severity is considered here equivalent to the difficulty index). As this question appears to be valid, we think that this incoherence is not due to chance but rather to a poor definition of the marking instructions, which misled the students in their assessments. However, with an internal consistency of the marks (Cronbach’s $\alpha$) equal to 0.61 (teacher’s view), we can consider this test as relevant for a peer assessment validity analysis.

Concerning the CA-2007 exam, the item analysis generated the results shown in Table 4. This table shows that the students marked the answers to question Q7 severely ($p=0.74$), whereas the teacher marked the answers to the same question with average severity ($p=0.59$). The difference is too large to be explained by chance alone. Another explanation is a poor indication of the marking instructions for this item. Indeed, in this question, the second and third parts of the answer expected are based on the result of the first part. This causes a problem with regard to the independence of the different parts of the marking scheme. The teacher’s grading takes into account the second and...
third ability even if the first one is not right. Since the marking instructions given to the students do not specify this, the students applied incoherent marking rules, different from those of the teacher. This automatically distorted our experiment, which was based on everyone using the same marking scale and marking rules. However, for the other questions in the test, we can consider that our experiment respected the hypothesis of clear marking scale and marking instructions. We will therefore disregard question “Q7” while performing the peer assessment validity analysis.

### Table 3. Item analysis of the CA-2006 exam

<table>
<thead>
<tr>
<th>Item</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>Mean : 8.5</td>
<td>SD : 2.8</td>
<td>α : 0.61</td>
<td>PM</td>
<td>Mean : 8.11</td>
<td>SD : 2.4</td>
<td>α : 0.55</td>
<td>SPM</td>
</tr>
<tr>
<td>Q1</td>
<td>0.12</td>
<td>-0.05</td>
<td>0.16</td>
<td>-0.06</td>
<td>0.15</td>
<td>-0.06</td>
<td>0.12</td>
<td>-0.03</td>
</tr>
<tr>
<td>Q2</td>
<td>0.59</td>
<td>0.15</td>
<td>0.52</td>
<td>0.10</td>
<td>0.52</td>
<td>0.10</td>
<td>0.49</td>
<td>0.12</td>
</tr>
<tr>
<td>Q3</td>
<td>0.16</td>
<td>0.13</td>
<td>0.17</td>
<td>0.10</td>
<td>0.16</td>
<td>0.10</td>
<td>0.10</td>
<td>0.24</td>
</tr>
<tr>
<td>Q4</td>
<td>0.86</td>
<td>0.25</td>
<td>0.82</td>
<td>0.21</td>
<td>0.81</td>
<td>0.21</td>
<td>0.74</td>
<td>0.17</td>
</tr>
<tr>
<td>Q5</td>
<td>0.64</td>
<td>0.45</td>
<td>0.66</td>
<td>0.51</td>
<td>0.66</td>
<td>0.51</td>
<td>0.62</td>
<td>0.47</td>
</tr>
<tr>
<td>Q6</td>
<td>0.71</td>
<td>0.22</td>
<td>0.68</td>
<td>0.12</td>
<td>0.68</td>
<td>0.12</td>
<td>0.74</td>
<td>0.19</td>
</tr>
<tr>
<td>Q7</td>
<td>0.84</td>
<td>0.22</td>
<td>0.81</td>
<td>0.28</td>
<td>0.80</td>
<td>0.28</td>
<td>0.80</td>
<td>0.21</td>
</tr>
<tr>
<td>Q8</td>
<td>0.88</td>
<td>0.02</td>
<td>0.80</td>
<td>0.07</td>
<td>0.77</td>
<td>0.07</td>
<td>0.68</td>
<td>0.17</td>
</tr>
<tr>
<td>Q9</td>
<td>0.71</td>
<td>0.39</td>
<td>0.76</td>
<td>0.32</td>
<td>0.75</td>
<td>0.32</td>
<td>0.72</td>
<td>0.37</td>
</tr>
<tr>
<td>Q10</td>
<td>0.67</td>
<td>-0.06</td>
<td>0.70</td>
<td>0.03</td>
<td>0.68</td>
<td>0.03</td>
<td>0.65</td>
<td>0.10</td>
</tr>
<tr>
<td>Q11</td>
<td>0.27</td>
<td>0.23</td>
<td>0.34</td>
<td>0.23</td>
<td>0.33</td>
<td>0.23</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>Q12</td>
<td>0.37</td>
<td>0.12</td>
<td>0.40</td>
<td>0.15</td>
<td>0.38</td>
<td>0.15</td>
<td>0.28</td>
<td>0.26</td>
</tr>
<tr>
<td>Q13</td>
<td>0.59</td>
<td>0.50</td>
<td>0.60</td>
<td>0.52</td>
<td>0.58</td>
<td>0.52</td>
<td>0.58</td>
<td>0.66</td>
</tr>
<tr>
<td>Q14</td>
<td>0.16</td>
<td>0.25</td>
<td>0.42</td>
<td>0.21</td>
<td>0.39</td>
<td>0.21</td>
<td>0.28</td>
<td>0.30</td>
</tr>
<tr>
<td>Q15</td>
<td>0.78</td>
<td>0.20</td>
<td>0.79</td>
<td>0.27</td>
<td>0.78</td>
<td>0.27</td>
<td>0.72</td>
<td>0.25</td>
</tr>
<tr>
<td>Q16</td>
<td>0.88</td>
<td>0.55</td>
<td>0.92</td>
<td>0.39</td>
<td>0.91</td>
<td>0.39</td>
<td>0.84</td>
<td>0.29</td>
</tr>
<tr>
<td>Q17</td>
<td>0.93</td>
<td>0.40</td>
<td>0.89</td>
<td>0.32</td>
<td>0.88</td>
<td>0.32</td>
<td>0.82</td>
<td>0.45</td>
</tr>
<tr>
<td>µ</td>
<td>0.60</td>
<td>0.23</td>
<td>0.61</td>
<td>0.22</td>
<td>0.60</td>
<td>0.22</td>
<td>0.55</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Correlation between severities (difficulty indexes): 0.97, 0.97, 0.97

Means differences: 0.39, 0.12, -0.83

Legend:
- p: Difficulty index
- r: Discrimination index
- α: Cronbach’s coefficient (reliability)
- µ: Means
- TM: Teacher marks
- PM: Peer marks
- SPM: Marks awarded by Peers and by self-assessment
- SM: Marks awarded by self-assessment
- NI: Number of items
- NS: Number of exam sheets
- SD: Standard deviation

### Table 4. Item analysis of the CA-2007 exam

<table>
<thead>
<tr>
<th>Item</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>NS=94</td>
<td>Mean : 11.03</td>
<td>SD : 3.00</td>
<td>α : 0.65</td>
<td>PM</td>
<td>NS=94</td>
<td>Mean : 10.50</td>
<td>SD : 2.60</td>
</tr>
<tr>
<td>Q1</td>
<td>0.24</td>
<td>0.29</td>
<td>0.26</td>
<td>0.21</td>
<td>0.25</td>
<td>0.23</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>Q2</td>
<td>0.04</td>
<td>0.13</td>
<td>0.06</td>
<td>0.16</td>
<td>0.05</td>
<td>0.14</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Q3</td>
<td>0.20</td>
<td>0.30</td>
<td>0.24</td>
<td>0.38</td>
<td>0.24</td>
<td>0.38</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>Q4</td>
<td>0.39</td>
<td>0.36</td>
<td>0.38</td>
<td>0.40</td>
<td>0.38</td>
<td>0.40</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Q5</td>
<td>0.64</td>
<td>0.44</td>
<td>0.66</td>
<td>0.38</td>
<td>0.64</td>
<td>0.40</td>
<td>0.56</td>
<td>0.34</td>
</tr>
<tr>
<td>Q6</td>
<td>0.48</td>
<td>0.16</td>
<td>0.48</td>
<td>0.28</td>
<td>0.47</td>
<td>0.28</td>
<td>0.45</td>
<td>0.17</td>
</tr>
<tr>
<td>Q7</td>
<td>0.59</td>
<td>0.24</td>
<td>0.74</td>
<td>0.34</td>
<td>0.72</td>
<td>0.34</td>
<td>0.63</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Correlation between severities (difficulty indexes): 0.97, 0.97, 0.97

Means differences: 0.39, 0.12, -0.83

Legend:
- p: Difficulty index
- r: Discrimination index
- α: Cronbach’s coefficient (reliability)
- µ: Means
- TM: Teacher marks
- PM: Peer marks
- SPM: Marks awarded by Peers and by self-assessment
- SM: Marks awarded by self-assessment
- NI: Number of items
- NS: Number of exam sheets
- SD: Standard deviation
As regards the GEE-2007 exam, the item analysis showed low reliability (0.41). Consequently, we could not use it to verify the peer assessment validity.

Checking the validity of peer assessment: Although the item analysis has already given us an idea of this validity, we preferred reinforcing our analysis with other metrics already used in other research dealing with the same issues as ours (Cho et al., 2006; Haaga, 1993; Mowl & Pain, 1995; Cheng & Warren, 1999; Stefani, 1994). These metrics (see Table 5) are Pearson’s correlation coefficient and the “T-Test”. We used these two metrics in the comparison between the teacher’s marks (TM) and those awarded by the peers (PM: peer marks disregarding self-assessment, SPM: peer marks incorporating the self-assessment, and SM: self-assessment mark). It should be noted that we did not take into account the GEE-2007 exam and the question Q7 of the CA-2007 exam, which were rejected by the item analysis.

The correlation coefficients between the teacher’s marks and those of the peers are high (0.90±2). The “T-Test”, applied to the marks awarded by the teacher and the marks awarded by peers, is positive. We can confirm that peer assessment is valid for the category of exams targeted in our research and that the assumptions outlined above are verified.

Table 5. T-test and correlation coefficient between the teacher marks and the peer marks.

<table>
<thead>
<tr>
<th>Number of papers</th>
<th>TM &amp; PM Comparison</th>
<th>TM &amp; SPM Comparison</th>
<th>TM &amp; PM Comparison</th>
<th>TM &amp; SPM Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable 1</td>
<td>TM</td>
<td>TM</td>
<td>TM</td>
<td>TM</td>
</tr>
<tr>
<td>Mean</td>
<td>8.426</td>
<td>8.426</td>
<td>9.806</td>
<td>9.806</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.830</td>
<td>2.830</td>
<td>2.651</td>
<td>2.651</td>
</tr>
<tr>
<td>Variable 2</td>
<td>PM</td>
<td>SPM</td>
<td>PM</td>
<td>SPM</td>
</tr>
<tr>
<td>Mean</td>
<td>8.1265</td>
<td>8.4009</td>
<td>9.7258</td>
<td>9.9566</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.3686</td>
<td>2.4082</td>
<td>2.3505</td>
<td>2.3331</td>
</tr>
</tbody>
</table>

T-Test Hypotheses: H0

- Signification level (%): 5%
- Number of degrees of freedom: 67
- T-Score found: 1.8529
- Critical value: 1.9944
- Conclusion: H0 accepted

Correlation coefficient (Pearson): 0.88

d of Cohen: 0.11

Effect size r: 0.06

Legend: TM: Teacher Marks
Our results compared to other research

Falchikov and Goldfinch (2000) performed a meta-analysis comparing the marks of peers to those of the teacher. This analysis does not deal with online peer assessment and mainly concerns undergraduate students. We have extracted some of the results from this meta-analysis by filtering the studies in relation to our research. These results prove that it is difficult to give an opinion on the validity of peer assessment while basing the research on previous studies. Indeed, the correlation coefficient between the peer marks and those of the teacher varies from 0.29 to 0.94. Peer assessment validity depends on several factors such as the reliability of the marking schemes and the clarity of the marking instructions. Our study proposes a method that makes it possible to avoid situations that could distort interpretation.

Falchikov and Goldfinch (2000) found that peer assessments resemble more closely teacher assessments when global judgments, based on well understood criteria, are used rather than when marking involves assessing several individual dimensions. Nevertheless, our study shows that it is possible, with an assessment involving several individual dimensions, to obtain a very good correlation between the peer marks and those of the teacher. These results, however, are limited to exams that mainly contain questions referring to exact science field and requesting simple calculations, some mathematical reasoning, short algorithms, and short texts.

Moreover, online peer assessment has been used during the past decade and some studies have dealt with the issue of validity (Cho et al., 2006; Tseng & Tsai, 2007). Our study confirms the results obtained by these studies. Indeed, Tseng et al carried out a study on the validity of peer assessment in a computer course involving 184 students. They claimed that the marks awarded by the peers largely corresponded to those awarded by the experts, thus indicating that peer assessment in high schools could be considered as a valid method of assessment. In addition, Cho et al (2006), while carrying out an experiment involving 708 students, proved that the overall assessments of at least 4 peers on a written examination are reliable and just as valid as the assessments made by the teacher. The results suggested that the teachers’ concerns regarding the reliability and validity of peer assessments should not prevent a peer assessment from being applied, at least with the appropriate scaffolding.

Conclusion

This study suggests trusting peer assessment when it is applied to exams that contain questions referring to the exact science field (calculations, mathematical reasoning, short algorithms and drafting of short texts) and when the exam paper is marked by at least four peers. The results also show that a combination of peer assessment with self-assessment gives a better validity to this assessment method. However, one must take some precautions (the quality of the marking schemes, the clarity of questions, the grading of the assessments quality, the technical assistance, and the assessment of examples).

The choice of using exams already carried out (in a traditional manner) in the classroom makes it possible to attain an authentic assessment and opens educational and administrative horizons (statistical treatment, item analysis, validity analysis, etc.). In particular, this method helps teachers identify what is wrong in their exams, thus giving them an opportunity to correct them. However, with respect to the teacher's workload, this choice is a disadvantage. Consequently, we must supplement our solution by creating a system that enhances online assessment by an automatic phase that makes it possible to put exams carried out in classroom online (this system is being developed).

Concerning the educational aspect, our research will be supplemented by an analysis of the interaction occurring between all of the actors (students and teachers) participating in the experiment. This will give us information on the impact that this assessment method has on the learning process.
Finally, regarding the scope of this research, our work will be supplemented by an experiment involving more students and courses. We will also explore the relationship that exists between peer assessment validity and variables like the number of assessors, the scale schemes, the student knowledge level, the knowledge fields to which the exams belong, and the cognitive levels required by the exam items.

References


Critical Thinking and ICT Integration in a Western Australian Secondary School

Graham McMahon
Office of Teaching and Learning, Curtin University of Technology, Perth Australia // G.McMahon@curtin.edu.au

ABSTRACT
This study examined the relationship between students working in a technology-rich environment and their development of higher order thinking skills. Based on a PhD thesis, which examined a greater range of relationships than can be reported here, this article focuses on developing critical thinking skills within a technology-rich environment. Staff and students from one school participated in the study. Data were collected to determine the degree of correlation between factors of the learning environment and the extent to which critical thinking skills were demonstrated by the students. Statistical correlations allowed relationships between environmental factors and critical thinking to be established. The results indicate that there are statistically significant correlations between studying within a technology-rich learning environment and the development of students’ critical thinking skills. Length of time spent in the environment has a positive, non-linear effect on the development of critical thinking skills. Students with better developed computing skills scored higher on critical thinking activities. This was most significant for students with better computer programming skills and the ability to competently manipulate Boolean logic. The research suggests that to develop students’ higher order thinking skills, schools should integrate technology across all of the learning areas. This will allow students to apply technology to the attainment of higher levels of cognition within specific contexts. This will need to be paralleled by providing students the opportunity to develop appropriate computer skills.

Keywords
Critical thinking, Higher order thinking skills (HOTS), Computer technology as a learning tool

Introduction
The International Conference on Computers in Education (ICCE, 2002) generated over 400 papers. Many of these detailed how ICT can be applied to higher order thinking skills, including Agnew (2002), Lee (2002) and Thomas (2002). Agnew (2002) iterated that researchers have tried to evaluate whether or not the use of technology has a major impact on student learning. Agnew (2002) continues by explaining that students are encouraged to develop higher order thinking skills, and that the results have been significant. However, no formal statement about how the achievement of these skills have been observed or measured has been made. While the application of notebook and/or hand-held computers within the learning environment has been examined at the tertiary level (Hewlett-Packard; Edith Cowen University, 2005) this area of research is in its infancy in secondary education, and little has been conducted within Western Australian secondary schools. Sherry and Jesse (2000) suggested that while educators seem to inherently ‘know’ that technology increases student achievement, measuring the increase is challenging. Branigan (2000, in Sherry & Jesse, 2000, p. 2) reported that in cases where teachers have applied a high degree of technology within the classroom, standardized test scores also were high. Trucano (2005) supports this argument, reporting that the application of Computer Aided Instruction (CAI) packages has resulted in improved test scores for mathematics and reading. However, he queries that these increased scores are indicative of a general increase in student learning. As schools have a far more important role than simply producing good test scores these results do not necessarily provide insight into overall student development. Trucano (2005) continues by arguing that, despite the results of many studies, the influence of ICT on student achievement is still difficult to measure.

This argument is reinforced by Sharma and Haigh (2008) who conducted a case study to examine how ICT is used to develop pedagogy. They concluded that the overall effect of ICT was difficult to isolate, due to the influence of other learning environment factors. Other studies have incorporated learning environment factors. Wanpen and Fisher (2004) used qualitative and quantitative techniques to assess the extent to which constructivism has been incorporated within computer classrooms, and to provide goals for further development. They concluded that it is possible to manipulate factors of the learning environment to make it more conducive for study in computer classes. Student reflections suggest that they found small group work helpful, demonstrating the value of a collaborative approach to learning.
Whilst examining the qualitative elements of a technology-rich environment, O’Dwyer, Russell, Bebell and Tucker-Seeley (2008) argued that traditional methods of assessing student performance may not be valid when technology is used. Existing instruments tend to measure critical and/or creative development in general terms, whereas technology is usually applied to specific learning areas (Russell, 2002, in O’Dwyer, Bebell, & Tucker-Seeley, 2008).

A similar argument was forwarded by McNabb, Valdez, Nowakowski, and Hawkes (2004). In their study of the application of technology to the development of mathematics and science education, the use of standardised testing is criticised. Instruments such as the California Achievement Test and Illinois Measure of Annual Growth Exam do not evaluate higher order thinking or technology skills, or the context in which these skills are developed.

To determine the relationships between ICT application, and other factors within the learning environment, and the development of higher order thinking skills, this paper examines the questions:

1. To what extent is computer-based technology integrated in the teaching-learning environment?
2. Is there a significant correlation between the students’ computer skills and the
devlopment of critical thinking skills?
   2.1. development of critical thinking skills?
   2.2. length of time spent within the environment?
   2.3. age of the technology?
3. Is there a significant correlation between the length of time spent within technology-rich environment and the development of critical higher order thinking skills?
4. Is there a significant correlation between the age of the technology and the development of critical higher order thinking skills?

For the purposes of this paper, higher order thinking skills are defined as those mental processes that allow students to develop factual, procedural, conceptual and metacognitive knowledge within the creative and critical domains described by Anderson and Krathwohl (2001).

Method

The approach taken by this research is that of a case study. Staff and students from one school participated in this case study. Data were collected by teachers as part of the normal teaching-learning program, supplemented by classroom observations and teacher interviews. The data were used to determine the degree of correlation between factors of the learning environment and the extent to which higher order thinking skills (HOTS) were demonstrated by the students. Collations of the statistically significant, and statistically insignificant, correlations allowed relationships between specific environmental factors and HOTS to be established.

The data source was the Year Nine student cohort in a metropolitan, independent girls’ school, providing a sample size of approximately 150 students. This school had been implementing a notebook computer program for nine years, in which all students in Years Five through to Year Ten use their notebooks across all learning areas, every day at school. The majority of students enrol in the junior years, although about one third of students begin in Years Seven and Eight, and approximately twelve students enrol in Year Nine. The Year Nine cohort constitutes students that have had between one and five years exposure to the technology-rich learning environment. Some students will buy a new notebook computer after three years; others will economise by keeping their original notebook computer for six years. Data about the length of enrolment at the school, and the age of their current notebook computer, are held within the school’s databases.

The instruments used were the
- Level of Technology Implementation (LoTI);
- Australian Schools Computer Skills Competition (ASCSC); and
- Ennis’ Weir Critical Thinking Essay Test (EWCTET).

The extent to which the learning environment can be considered as technology-rich was determined by implementing the LoTI process (Moersch, 1999). The LoTI instrument is used to calculate the level of computer efficiency of an institution. It draws on quantifiable data obtained through observations and structured interviews (Moersch, 2003). The different levels are summarised in Table 1.
Table 1: Description of LoTI Levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Pedagogy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Little or no computer use</td>
</tr>
<tr>
<td>1</td>
<td>Technology-based instruction is predominately about the technology</td>
</tr>
<tr>
<td>2</td>
<td>Technology use is peripheral or dispensable</td>
</tr>
<tr>
<td>3</td>
<td>Technology use is adapted to fit with traditional goals</td>
</tr>
<tr>
<td>4</td>
<td>Technology is used for collaborative instruction</td>
</tr>
<tr>
<td>5</td>
<td>Technology supports self directed, collaborative learning</td>
</tr>
<tr>
<td>6</td>
<td>Technology supports self directed, collaborative learning; students &amp; teachers are learners and researchers</td>
</tr>
</tbody>
</table>

The technology skills of students were determined from the externally marked ASCSC. The ASCSC is a national test that students sit annually. The competition is designed to assess students’ computer skills and knowledge outlined in the various State curricula. The questions reflect the importance of the higher order thinking skills, and present an opportunity for students to perform a range of tasks using these skills. Measures of the ASCSC produced Cronbach’s alpha coefficients ranging from 0.55 through to 0.70 (McMahon, 2007), which suggests that the ASCSC (2005) provides a reliable means of measuring students’ computer skills.

The EWCTET requires student to formulate a complex argument in response to an earlier argument. In doing so the test measures the student’s ability to analyse an argument and provide a logical, coherent response, including creative and critical dimensions (Ennis & Weir, 1985). In a later paper (Ennis, 1998) the test is reported as being valid and reliable, with inter-rater reliabilities in the order of 0.72 – 0.93.

Data concerning students’ length of enrolment at the school, and the age of their computers, were extracted from the school’s databases. The ASCSC and EWCTET were marked according to predetermined answers keys. Whilst the ASCS marking key is prescriptive the EWCTET allows for some teacher judgment. All of these instruments provide a raw numerical score. Data analysis involved the correlation of results to determine the relationships that exist between the different learning environment factors.

Questions two, three and four were framed as a series of hypotheses that could be accepted or rejected based on statistical analysis:

1.1 \( H_0: \) There is no significant correlation between students’ computer skills and the development of critical thinking skills.
\( H_1: \) There is a significant correlation between students’ computer skills and the development of critical thinking skills.

1.2 \( H_0: \) There is no significant correlation between students’ computer skills and the length of time spent within the environment.
\( H_1: \) There is a significant correlation between students’ computer skills and the length of time spent within the environment.

1.3 \( H_0: \) There is no significant correlation between students’ computer skills and the age of the technology.
\( H_1: \) There is a significant correlation between students’ computer skills and the age of the technology.

2 \( H_0: \) There is no significant correlation between the length of time spent within technology-rich environments and the development of critical thinking skills.
\( H_1: \) There is a significant correlation between the length of time spent within technology-rich environments and the development of critical thinking skills.

3 \( H_0: \) There is no significant correlation between the age of the technology and the development of critical thinking skills.
\( H_1: \) There is a significant correlation between the age of the technology and the development of critical thinking skills.

For the purposes of testing the hypotheses, correlation coefficients between data gained from these instruments were calculated. The correlations were determined by calculating Spearman’s correlation coefficients relevant to each
research question. ASCSC scores were recorded as student percentile rankings; for consistency between the sets of correlations, Spearman’s rho was calculated in place of Pearson’s r. Data from the instruments were reinforced with data collected from staff and student interviews, and classroom observations. Concurrently, the LoTI instrument (Moersch, 1999) was used to quantify the technology use within the school and therefore determine the level at which the school is operating, with respect to technology integration.

**Results**

**Technology Integration**

The data in Tables 2 and 3 indicate that technology is predominantly used at Levels Four and Five, as measured by the LoTI scale, which suggests an emphasis on lessons that implement technology used across the learning areas to address higher orders of learning.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Level</th>
<th>Learner Use %</th>
<th>No. of Computers</th>
<th>Product (B x C x D) / 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>Non-use**</td>
<td>0</td>
<td>33.33</td>
<td>291</td>
<td>0.00</td>
</tr>
<tr>
<td>Awareness</td>
<td>1</td>
<td>4.87</td>
<td>291</td>
<td>14.17</td>
</tr>
<tr>
<td>Exploration</td>
<td>2</td>
<td>8.53</td>
<td>291</td>
<td>49.64</td>
</tr>
<tr>
<td>Infusion</td>
<td>3</td>
<td>15.13</td>
<td>291</td>
<td>132.08</td>
</tr>
<tr>
<td>Integration</td>
<td>4</td>
<td>18.53</td>
<td>291</td>
<td>215.69</td>
</tr>
<tr>
<td>Expansion</td>
<td>5</td>
<td>12.07</td>
<td>291</td>
<td>175.62</td>
</tr>
<tr>
<td>Refinement</td>
<td>6</td>
<td>7.54</td>
<td>291</td>
<td>131.65</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>291</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>718.85 / (291 x 4) = 0.62</td>
<td></td>
<td></td>
<td>Computer Efficiency Rating = 62%</td>
</tr>
</tbody>
</table>

*Data were collected by an anonymous survey.

**Non-use includes time spent on classroom administration (E.G. roll call), introducing the lessons, giving instructions to students and concluding lessons.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Level</th>
<th>Learner Use %</th>
<th>No. of Computers</th>
<th>Product (B x C x D) / 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>Non-use**</td>
<td>0</td>
<td>33.33</td>
<td>25</td>
<td>0.00</td>
</tr>
<tr>
<td>Awareness</td>
<td>1</td>
<td>0.77</td>
<td>25</td>
<td>0.19</td>
</tr>
<tr>
<td>Exploration</td>
<td>2</td>
<td>3.83</td>
<td>25</td>
<td>1.92</td>
</tr>
<tr>
<td>Infusion</td>
<td>3</td>
<td>13.79</td>
<td>25</td>
<td>10.34</td>
</tr>
<tr>
<td>Integration</td>
<td>4</td>
<td>28.35</td>
<td>25</td>
<td>28.35</td>
</tr>
<tr>
<td>Expansion</td>
<td>5</td>
<td>8.43</td>
<td>25</td>
<td>10.54</td>
</tr>
<tr>
<td>Refinement</td>
<td>6</td>
<td>11.50</td>
<td>25</td>
<td>17.25</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>68.59 / (25 x 4) = 0.68</td>
<td></td>
<td></td>
<td>Computer Efficiency Rating = 68%</td>
</tr>
</tbody>
</table>

* Data were collected through direct classroom observation.

**Non-use includes time spent on classroom administration (E.G. roll call), introducing the lessons, giving instructions to students and concluding lessons.

Assessing the level of technology integration is not an exact science. Classroom observations require the approval of the teacher. Within an environment that promotes technology, there could be some reluctance to allow observations in a classroom that may not promote technology integration. Those teachers that readily accept observations of their lessons are likely to be those who confidently and gainfully employ technology in their classes.
To address this predicament, data for the LoTI evaluation were collected by two methods. Data in Table 2 was collected by an anonymous survey, in which teaching staff indicated the types of activities that were occurring in their classes. Data in Table 3 was collected through direct classroom observations across eight learning areas. LoTI efficiency ratings were then calculated based on both sets of data. While there is no statistical test that can be used to determine the significance of this data, the two calculated efficiency ratings of 68% and 62%, shown in Tables 2 and 3 are close, which suggests that these approaches are reliable. The efficiency rating between 60% and 70% indicates a high level of computer integration across the learning areas.

Computer Skills

Table 4 shows that, on average, the research cohort scores are 2% higher than the State mean scores. Furthermore, the research cohort has greater than 40% of its students in the top 30% of the State scores.

<table>
<thead>
<tr>
<th>Year</th>
<th>School Mean Score ($\bar{x}/45$)</th>
<th>State Mean Score ($\bar{x}/45$)</th>
<th>% of School Cohort in the State:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>50%</td>
</tr>
<tr>
<td>2005</td>
<td>Mean Score</td>
<td>24.1</td>
<td>58.9</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>404</td>
<td>40.8</td>
</tr>
<tr>
<td>2006</td>
<td>Mean Score</td>
<td>20.0</td>
<td>66.3</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>380</td>
<td>40.8</td>
</tr>
</tbody>
</table>

Prior to 2005 students entered this competition on a voluntary basis, resulting in low numbers relative to later years. Student performance for 2002-03 compared to the State are summarised in Table 5.

<table>
<thead>
<tr>
<th>Year</th>
<th>School Mean Score ($\bar{x}/45$)</th>
<th>State Mean Score ($\bar{x}/45$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>Mean Score</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>19</td>
</tr>
<tr>
<td>2003</td>
<td>Mean Score</td>
<td>29.7</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>56</td>
</tr>
</tbody>
</table>

Whilst Table 5 indicates a higher mean result for the school than for the State, this may not be a true reflection of the students’ level of computing skills. As the test was sat by volunteers, it could be argued that the volunteers had an ‘above average’ interest in technology, which could inflate the school’s results. However, this could also apply to the students of other schools involved in the test.

The 2% difference between the mean results for the school and State has been maintained when the complete cohort is considered in Table 4. Furthermore, 40% of the school cohort results lie within the top 30% of the State results. Eleven percent of the school cohort lies within the top 10% of the State results. This can be further analysed using the Chi-squared test of the hypotheses:

\[ H_0: \text{There is no significant difference between the observed and expected mean values.} \]
\[ H_1: \text{There is a significant difference between the observed and expected mean values.} \]

Chi-squared values are shown in Table 6.

<table>
<thead>
<tr>
<th>Observed % of School Cohort in the State:</th>
<th>50%</th>
<th>30%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>66.3</td>
<td>40.8</td>
<td>11.6</td>
</tr>
<tr>
<td>Expected</td>
<td>50.0</td>
<td>30.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 9.491^{**} \]

** Significant at the 0.01 level
The calculated Chi-squared value in Table 6 indicates that the null hypothesis ($H_0$) should be rejected, in favour of the alternate hypothesis ($H_1$), that there is a significant difference between the observed and expected mean values. The scope of technology integration in the ‘State’ cohort is unknown. The effect of cultural differences between the environments that may affect the test results are not reported here. Those who are interested are directed to http://adt.curtin.edu.au/theses/available/adt-WCU20070525.172626/.

Critical Thinking Skills

Three groups of students were video recorded as they completed the EWCTET, to identify the different levels of cognitive activity described by Anderson and Krathwohl, (2001) that they employed. Cognitive activity was determined from the questions asked of the teacher, discussions between students, and their behaviour during the task. Examples are shown in Table 7.

<table>
<thead>
<tr>
<th>Cognitive Dimension</th>
<th>Knowledge Dimension</th>
<th>Example of student dialogue and/or activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disengaged*</td>
<td>Factual</td>
<td>[Discussion of TV shows]</td>
</tr>
<tr>
<td></td>
<td>Remembering</td>
<td>I’m not sure what we have to do</td>
</tr>
<tr>
<td></td>
<td></td>
<td>We have to read the letter</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td>Do you choose one of them [Tasks]??</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>No, you have to do both.</td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>-</td>
</tr>
<tr>
<td>Understanding</td>
<td>Factual</td>
<td>[Listening to instructions]</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td>Are we supposed to put down our own objections and things?</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>So, we’re responding as though this letter was written to us.</td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>-</td>
</tr>
<tr>
<td>Applying</td>
<td>Factual</td>
<td>[Working on set task following prescribed format – identifying, clarifying and responding to a point of view]</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>-</td>
</tr>
<tr>
<td>Analysing</td>
<td>Factual</td>
<td>From [paragraph] five onwards it’s all about accidents.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>He’s using a lot of facts, but it’s too much facts because it’s kind of the same thing</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td>I don’t get what it means by … [Reads aloud from task]</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>I disagree with this point here. [Points to screen]</td>
</tr>
<tr>
<td>Evaluating</td>
<td>Factual</td>
<td>It’s not very clear. [The argument]</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td>Explains personal response to presented arguments to partner, relating to authentic conditions close to the school</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>The debate is not good – he needs to get proof</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The person needs to write the letter more clearly so that whom he is sending it to can understand it better.</td>
</tr>
<tr>
<td>Creating</td>
<td>Factual</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Conceptual</td>
<td>Make it that they have to pay for it, or get a parking ticket.</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>Maybe they could widen the streets by taking out the pathways. [Elaborates in detail, based on personal knowledge of a similar situation in rural Australia]</td>
</tr>
<tr>
<td></td>
<td>Metacognitive</td>
<td>-</td>
</tr>
</tbody>
</table>

*Disengaged activities are those that occur when students engage in activities not related to the assignment

The time each group of students spent operating at the different cognitive levels is shown in Table 8.

<table>
<thead>
<tr>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>% Lesson time at each level</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

274
Students’ performances for this task are shown in Table 9.

<table>
<thead>
<tr>
<th>Critical Thinking Skill</th>
<th>Possible</th>
<th>Maximum Gained</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognise False Analogy</td>
<td>3.00</td>
<td>3.00</td>
<td>0.45</td>
</tr>
<tr>
<td>Recognise Irrelevant Argument</td>
<td>3.00</td>
<td>3.00</td>
<td>0.58</td>
</tr>
<tr>
<td>Recognise Correct Argument</td>
<td>6.00</td>
<td>6.00</td>
<td>2.33</td>
</tr>
<tr>
<td>Recognise No Support</td>
<td>3.00</td>
<td>3.00</td>
<td>0.71</td>
</tr>
<tr>
<td>Recognise Alternative Arguments</td>
<td>3.00</td>
<td>3.00</td>
<td>0.79</td>
</tr>
<tr>
<td>Recognise Poor Arguments</td>
<td>3.00</td>
<td>3.00</td>
<td>0.86</td>
</tr>
<tr>
<td>Recognise Poor Experimental Design</td>
<td>3.00</td>
<td>3.00</td>
<td>0.75</td>
</tr>
<tr>
<td>Recognise Poor Definition</td>
<td>5.00</td>
<td>5.00</td>
<td>0.54</td>
</tr>
<tr>
<td>Recognise Emotive Language &amp; Provide Coherent Summary</td>
<td>2.00</td>
<td>2.00</td>
<td><strong>5.96</strong></td>
</tr>
<tr>
<td>Total</td>
<td>29.00</td>
<td>29.00</td>
<td>5.96</td>
</tr>
</tbody>
</table>

* This is the highest score attained by a student, not the sum of the maximum scores gained by the student cohort  
  ** This is the mean total score, not the sum of the mean scores for each question

When addressing the issues raised in the EWCTET students did not automatically move to a level of higher cognition. Recorded observations between students indicate that they move to higher order thinking from a basis of lower order thinking during the early stages of the lesson. All observed groups spent time at the lower taxonomic levels of Anderson and Krathwohl (2001), and the interactions between the students allowed them to achieve higher order thinking skills. It was not always possible to clearly identify the dimension into which student dialogue or behaviour belonged. This resulted in some of the dimensions be merged and others apparently unaddressed.

Group One spent little time in studying the arguments to be addressed. Instead, they launched into their response at a low cognitive level and spent most of their time engaged at the applying level, frequently reverting to off-task behaviour. This group was engaged at an analysis level, the lowest of the higher cognitive levels, for less than five per cent of the class time, and did not reach a higher level of engagement.

The short bursts of time spent at the applying and analysis levels were quickly followed by off-task activities. This group would often ask members within the group questions such as “What does this mean?” or “What do we do here?” indicating a poor understanding of the task.

Group Two followed a similar pattern of off-task and low-cognitive activity. The first 17 minutes of the lesson were spent discussing the requirements of the task at remembering and understanding levels, alternating with off-task behaviour. It is probable that this would have continued for a greater proportion of the lesson had not the teacher intervened and re-focussed their activities. However, this group continued to struggle with the task, and at no stage were they observed to be operating at a level higher than applying.
Group Three displayed a learning pattern that suggests better personal organisation, better understanding of the task, and the ability to operate at higher cognitive levels. The early stages of the lesson were characterised by quiet reflection of the task, followed by group discussion in which the requirements were clarified. This was followed by discussions indicative of higher order thinking as described in Table 7. After higher level discussions the group appears to move to a lower cognitive level (applying) during which they are using their discussion to form their critical thinking responses.

Data in Table 9 initially suggest that the general level of critical thinking skills is low. An average score of $\frac{9}{29}$ is considerably short of the generally accepted ‘pass mark’ of 50%, or $\frac{15}{29}$. It should be remembered that the EWCTET is designed for secondary and tertiary students (Ennis & Weir, 1985). It is reasonable to assume that tertiary students would, on average, perform better than secondary students. Therefore an acceptable pass mark for tertiary students would be higher than that for those still at a junior grade, secondary level.

Rather than considering the test in terms of pass or fail, it can be viewed as an open-ended instrument. Data from previous administrations of the EWCTET is shown in Table 10 (Ennis, 1998).

<table>
<thead>
<tr>
<th>Source</th>
<th>Age Group</th>
<th>Number of Students</th>
<th>Mean Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davidson &amp; Dunham</td>
<td>1st Year, Junior College</td>
<td>19</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>4.7</td>
</tr>
<tr>
<td>Hatcher (2004)</td>
<td>1st Year University</td>
<td>169</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>119</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>178</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>178</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>164</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>169</td>
<td>6.9</td>
</tr>
<tr>
<td>Hatcher (1995)</td>
<td>Final Year University</td>
<td>66</td>
<td>9.6</td>
</tr>
</tbody>
</table>

If the data Table 10 are compared with the data from Table 9 (mean score = 5.96) it appears that the general standard of critical thinking skills transcends the critical thinking skills of college level students. When examining the thinking skills measured by the EWCTET, the highest mean scores relate to the students’ ability to recognise a logical argument and arguments involving experimental design. Whilst critical thinking skills are not specifically taught within this learning environment, all students were in their third year of science classes that do address the scientific method, which includes the accurate design and execution of experiments.

Within the researched learning environment, science classes use technology extensively. Spiro’s cognitive flexibility theory (Spiro, Feltovich, Jacobson & Coulson, 1991), explains how students can restructure their thinking as a means of adapting to new situations (Swindler, 2001). Swain, Greer and van Hover (2001) argue that the flexibility provided by computers make them an ideal learning tool employing the cognitive flexibility theory. It seems evident that students are transferring their critical thinking skills, developed within a technology-based science education environment, to other learning situations.

Data in Tables 4 and 5 indicate that above average computer skills exist within this learning environment when compared to those of the State. Similarly, data in Tables 9 and 10 indicate above average critical thinking skills within this learning environment. Spearman’s Correlations between the factors within the technology-rich learning environment are shown in Table 11.

<table>
<thead>
<tr>
<th></th>
<th>EWCTET</th>
<th>Length of Enrolment</th>
<th>Age of Computers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCSC Rho</td>
<td>0.307**</td>
<td>0.154*</td>
<td>0.039</td>
</tr>
<tr>
<td>N</td>
<td>116</td>
<td>135</td>
<td>133</td>
</tr>
<tr>
<td>Length of Enrolment Rho</td>
<td>0.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>131</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Reject the null hypothesis (1.1: H₀). Accept the alternate hypothesis (1.1: H₁) that there is a significant correlation between students’ computer skills and students’ critical thinking skills.

Reject the null hypothesis (1.2: H₀). Accept the alternate hypothesis (1.2: H₁) that there is a significant correlation between students’ computer skills and their length of time within the environment.

Accept the null hypothesis (1.3: H₀). There is no significant correlation between the age of the students’ computers and the development of students’ computing skills.

Accept the null hypothesis (2: H₀). There is no significant correlation between students’ critical thinking skills and their length of time within the environment.

Accept the null hypothesis (3: H₀). There is no significant correlation between the age of the students’ computers and the development of students’ critical thinking skills.

The EWCTET examines different aspects of critical thinking. This test cannot be reliably broken into these aspects but it can be treated as a reliable measure of critical thinking in total (Ennis & Weir, 1985; McMahon, 2007). The ASCSC is designed with questions that specifically address different scales of computer skills, and these scales have been shown to be reliable by McMahon (2007). Correlations between the score for these computer skill scales and scores for critical thinking are shown in Table 12.

<table>
<thead>
<tr>
<th>ASCSC scales</th>
<th>EWCTET Rho</th>
<th>Productivity Tools</th>
<th>Internet</th>
<th>Communication</th>
<th>Hardware</th>
<th>Software</th>
<th>Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.108</td>
<td>0.229**</td>
<td>0.231**</td>
<td>0.034</td>
<td>0.196*</td>
<td>0.212*</td>
</tr>
<tr>
<td>N</td>
<td>116</td>
<td>117</td>
<td>116</td>
<td>117</td>
<td>116</td>
<td>117</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (1-tailed)
**Correlation is significant at the 0.01 level (1-tailed)

The data show that the correlations are not significant for all of the scales used to measure students’ computer skills. The greatest significance occurs for those scales that address higher order thinking skills, for example, the Internet and Programming scales. These scales require students to be able to apply Boolean logic, and analyse and evaluate programming script.

In the researched learning environment there are intakes of new students each year. By dividing the Year Nine cohort into series of groups, statistical analysis between the series of groups is possible. The cohort was divided into students with less than one year in the environment and greater than one year; less than two years, greater than two years; and so on up to less than nine years, greater than nine years. Statistical correlations for these groups are shown in Table 13 and Figure 1.

<table>
<thead>
<tr>
<th>Length of Enrolment (Years)</th>
<th>EWCTET</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;=1</td>
<td>0.080</td>
</tr>
<tr>
<td>&gt;=2</td>
<td>-0.115</td>
</tr>
<tr>
<td>&gt;=3</td>
<td>0.114</td>
</tr>
<tr>
<td>&gt;=4</td>
<td>0.098</td>
</tr>
<tr>
<td>&gt;=5</td>
<td>0.157*</td>
</tr>
<tr>
<td>&gt;=6</td>
<td>0.157*</td>
</tr>
<tr>
<td>&gt;=7</td>
<td></td>
</tr>
<tr>
<td>&gt;=8</td>
<td></td>
</tr>
<tr>
<td>&gt;=9</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 provides a graphical representation of these correlations, from which a second order polynomial line-of-best-fit can be inferred. This describes the general trend between the length of time spent in the environment and the development of critical thinking skills.

Discussion

Papert (1972) imagined an educational environment in which students used computers as an extension of their mind. An advocate of teaching children how to program computers with Logo, Papert believed that the logical nature of programming coupled with the imaginations of children allows them to understand abstract principles by addressing them within a concrete framework. The data in Table 12 support this belief; students with better developed programming skills, developed within a technology-rich environment, score higher on critical thinking exercises. While this is not to state that the better programming skills cause higher order thinking, it should be noted that the computer skills are explicitly taught within this educational environment. Higher order thinking skills are not.

The data show a significant statistical correlation between the length of time students spend in the technology-rich environment and the development of their computer skills. It could be expected that, given this correlation, there would also be a significant statistical correlation between the students’ computer skills and the age of the computers.

Paradoxically, it could be argued that older computers are less reliable or efficient than more recent models, which could have a negative effect on the students’ learning and development of computer skills, yet there is no significant statistical correlation, positive or negative, between students’ computer skills and the age of the technology.

An important component of this technology-rich learning environment is the comprehensive technology support offered, in terms of hardware maintenance and software updates. Irrespective of the age of the notebook computers, the turn-around time for repairs is low, most being returned to the students on the same day. When a student’s
computer is being repaired the student is immediately given a replacement, in order that she may continue her lessons with minimal disruption.

One function of the Learning Technologies department within this school is to actively search for software to assist students’ learning. Any potential package is then tested on the different notebook models to ensure that is will perform adequately. When the software updates and hardware maintenance are considered, the insignificant statistical correlation between the age of the computer and the development of higher order thinking skills is understandable. Certainly, there may be some angst from students enviously regarding their fellow students’ latest acquisition, but this effect is minimal.

Teaching specific skills related to computer programming enhances students’ attainment of higher order thinking skills. The programming skills include developing an understanding of Boolean logic, top-down approaches to solving problems and exploring data manipulation from novel dimensions. The development of computer skills within a socially constructivist, technology-rich environment allows students to attain higher order thinking skills on par with tertiary level students.

Schools should endeavour to integrate technology across all of the learning areas. This will allow students to apply technology to the attainment of higher levels of cognition within specific contexts. If students are to apply computer-based technology to their studies they must be given the opportunity to develop appropriate computer skills. As demonstrated by the data, this does not necessarily require the latest model computers. The power of computers in education lies in embedding them in the curriculum and not using them as embellishments.

O’Dwyer, Russell, Bebell and Tucker-Seeley (2008), and McNabb et al. (2004), contended that traditional methods of assessing student performance measured higher order thinking in general terms rather than in specific learning areas to which technology is applied. Within this study groups students used technology extensively within the English learning area, and it is within this area that the EWCTET measures critical thinking.

There is no simple linear correlation between the length of enrolment in a technology-rich environment and the development of critical thinking skills. Closer analysis of the data provides more illumination. Table 13 shows that there is a significant statistical difference in the critical thinking skills (as measured by the EWCTET) of students who have been immersed in a technology-rich culture more than five years when compared to those who have been a part of the same culture for less than five years. This difference is maintained for students who have been a part of the culture for more than six, seven, eight and nine years. The cultural experiences from pre-school and/or kindergarten are not statistically significant with respect to critical thinking in Year Nine.

The data shown in Table 13 and Figure 1 suggest that there is a negative influence of the socio-cultural environment on the development of critical thinking for students with less than two years immersion in the environment. Adeyemo (2005) describes how a transition into secondary school is associated by a decline in academic achievement. This is not a new phenomenon, having been previously reported by Maehr (1974).

Maehr describes how our social experiences affect the perception of our environment and therefore affects the acquisition of knowledge. As the social group to which a student belongs is changed, achievement also changes. In a more recent study, Tonkin and Watt (2003) have associated the decline in academic performance with both the transition to a new environment and the onset of adolescence. Given that all of the students in this research cohort are of the same age group, the adolescent factor is negated. There is greater influence from moving into a new environment. Students moving into this technology-rich environment have an added burden related to the use of technology in all of their classes. Being unfamiliar with the protocol of electronically delivered learning materials, and even unfamiliarity of the technology itself, initially detract from their ability to achieve higher cognition with the technology.

At this point a caveat needs to be added. The population sample sizes used in this research were around 150, which may be considered large. Spencer (1995) argued that statistically significant correlations derived from large samples do not necessarily have a practical value. If the results are statistically significant, but small, then the cost of implementing the results in a social setting may not be justifiable (Morgan, Leech, Gloeckner & Barrett, 2006). Cohen (1992) suggested that the practical significance of statistical correlations with values of 0.1, 0.3 and 0.5 should be considered small, medium and large respectively. Accordingly, the statistically significant correlations presented in this paper would be regarded as having a small to medium practical significance. Schools considering
implementing a school-wide computer integration program would need to closely examine the associated costs with the perceived benefits. The intention of the Australian Labor Party to provide a laptop for every student (Winterford, 2007) implies that such programs are justified, although the debate is not yet over (Parker, 2008).

Conclusions

There is a statistically significant correlation between students’ computers skills measured by the ASCSC and their level of critical thinking skills measured by the EWCTET. The correlation exists between critical thinking skills and higher order computer skills, represented by the ASCSC scales of Internet and Programming. Students with better computing skills demonstrate a higher level of critical thinking.

These results address the concerns of Agnew (2002) and Trucano (2005) by providing data that links the use of technology to student learning. The instruments used – the ASCSC and the EWCTET – have been demonstrated as a reliable and valid method of measuring this relationship.

There is a statistically significant, positive correlation between the length of time that students spend in the technology-rich environment and their development of computer skills. Regular updating of software, and a well-established hardware maintenance program, mean that there is no statistically significant correlation between the age of the technology and the development of computer skills. The use of applications packages allows students to quickly and accurately process data, and therefore create knowledge. Time saved permits more time to be spent exploring specific subjects at deeper levels; exploration which is further enhanced by using the technology.

The statistically significant correlations are low and/or medium, as described by Cohen (1992). Further research is necessary to establish the practical significance of these findings.

While technology skills include the lower level manipulation of applications programs, they include developing higher order skills such as computer programming. Using productivity tools effectively allows subjects to be explored in greater detail. Computer programming develops skills in logic, critical thinking and creative thinking. This research indicates that the technology has a greater influence on critical thinking. Application of computing to assist the development of creative thinking needs to be further developed. Consideration should also be given to the establishment of technology-based teaching-learning programs to develop generic higher order thinking skills.

Acknowledgements

Educational Assessment Australia Management Committee, University of New South Wales for permission to incorporate the ASCSC into this study. Graham Rixon, Principal, Penrhos College, for permission to conduct the research.

References


Edith Cowan University (2005). Portable computing is catching on. ECUUpdate, 3, 3.


281
Learner and Instructional Factors Influencing Learning Outcomes within a Blended Learning Environment

Doo Hun Lim¹ and Michael Lane Morris²

¹Adult and Higher Education, University of Oklahoma, USA // Tel: +1 405-325-7941 // Fax: +1 405-325-2403 // dhlim@ou.edu
²University of Tennessee, Department of Management, USA // Tel: +1 865-974-6291 // Fax: +1 865-974-2048 // mmorris1@utk.edu

ABSTRACT
Among the many studies focusing on the effect of learner and instructional variables on learning outcomes, few studies have investigated the effect of these variables and their mediating mechanisms influencing students’ learning within a blended learning environment. This study examined the influence of instructional and learner variables on learning outcomes for a blended instruction course offered for undergraduate students. Data analysis indicated that age, prior experiences with distance learning opportunities, preference in delivery format, and average study time are those learner antecedents differentiating learning outcomes among groups of college students. From a regression analysis, the influence of learner, instructional, and motivational variables on learning outcomes found to be consolidated around one variable in learning application.

Keywords
Blended instruction, Instructional variables, Learner variables, Learning motivation

Introduction
As a result of the advancement in communication and network technologies, more innovative instructional delivery and learning solutions have emerged in order to provide meaningful learning experiences for learners in academic settings. Blended instruction is one of the various methods being used to deliver meaningful learning experiences. The use of blended instruction is growing rapidly because instructors believe diverse delivery methods may significantly enhance learning outcomes as well as increase student satisfaction from the learning experience. Previous research has indicated learning quantity and quality suffers when learners are solely and completely immersed in technology-based instructional delivery methods. Reasons that help explain this learning decline include a lack of human interaction (Laurillard, 1993), acute learning curves in adjusting to the new technology tools, delayed feedback, procrastination in learning (Lim, 2002), and lower motivation to read online learning materials (Lim & Kim, 2003). In order to eliminate the deleterious effects associated with the sole use of technology-based learning methods, many colleges and universities have adopted various forms of blended instruction to more effectively deliver instructional content to students and promote their learning.

Research has suggested a number of reasons in support of using blended instruction. Osguthorpe and Graham (2003) found that blended instruction methods improved pedagogy, increased access to knowledge, fostered social interaction, increased the amount of teacher presence during learning, improved cost effectiveness, and enhanced ease of revision. Similarly, Chung and Davis (1995) reported that blended instruction provided learners with greater control over the pace of learning, instructional flow, selection of resources, and time management. In studies involving self-regulated and interpersonal learners, blended instruction was found to be effective in addressing diverse learning styles (Bielawski & Metcalf, 2003). While these findings support the positive effect of blended instruction on individual learner’s learning, there is a lack of research to examine what learner and instructional variables within blended learning environment individually or collectively influence student learning especially in academic settings.

A lingering research question in college education is determining if traditional instructional methodologies satisfy high level learning needs like having opportunities to apply, synthesize, and integrate students’ learning experiences throughout their learning processes as well as low level learning needs like the acquisition of information, knowledge, and skills. Previous research suggests that reinforcement of learning is amplified when students are given in-class opportunities to practice and apply what they have learned, or are encouraged to immediately transfer their learning upon returning to their jobs and tasks (Lim & Johnson, 2002; Sullivan, 2002). In this study, the phrase ‘application of learning’ refers to the degree to which students transfer, use, and apply learned knowledge and skills to their current studies or to current jobs and tasks. Given the growing accountability interest among instructors and...
stakeholders of college education for identifying and demonstrating tangible evidence of instructional effectiveness in college education, the assessment of learning application outcomes has become a critical issue for researchers in education discipline. This study seeks to add to this growing body of literature.

In this study, we are seeking to identify what instructional factors, learner variables, and learning motivation and involvement factors within a blended learning environment influence a group of undergraduate students’ learning and learning application for an upper level undergraduate course conducted at a southeastern university. The following research questions were developed to address our research purpose.

1. What instructional factors, learner variables, and learning motivation and involvement factors within a blended learning environment influence students’ learning and learning application?
2. What are the reasons given for the students’ perceived high or low learning and learning application?

Theoretical Background

Blended Instruction

Several definitions of blended instruction exist. One frequently cited definition of blended instruction is the mix of traditional and interactive-rich forms of classroom instruction with learning technologies (Bielawski & Metcalf, 2003) Other researchers defined blended instruction as an integrated approach utilizing strategically planned instructional or non-instructional approaches for fostering learning and performance (Rossett, Douglas, & Frazee, 2003). In this study, blended instruction is defined as the appropriate mix and use of face-to-face instructional methods and various learning technologies to support planned learning and foster subsequent learning outcomes.

In determining the appropriate mix of blended instruction, Driscoll (2002) provides a framework for the possible blends of instructional delivery methods consisting of Web/Computer-Based Training, Web/Electronic Performance Support System, Web/Virtual Asynchronous Classrooms, and Web/Virtual Synchronous Classrooms. While Driscoll’s framework of blended instruction is widely used among instructional designers to develop blended instruction, a limitation of the framework is that blending is based on the mix between web-based technologies with other kinds of technology-based instructions and lacks inclusion of diverse instructional methods used in classroom settings (i.e., case studies, experiential learning, problem based learning, etc.). In determining the appropriate mix of instructional methods for blended instruction, one should consider several facets of learning modes such as live events, self-paced learning, collaboration, assessment, and learning support materials that facilitate students’ learning during the process of blended instruction (Carmen, 2002).

Learner Variables

Research studies involving student’s learning have focused on the correlations of antecedent variables like learner’s demographic variables with course participation, learning increase, and/or satisfaction. To illustrate, regarding gender, female students have self-reported experiencing more connection toward the learner and the social community in online learning environments (Rovai & Baker, 2005) and expressed more positive attitudes for online learning experiences (Sullivan, 2001). Similarly, male students used more formal and lengthier messages in online discussions when compared to female students (Taplin & Jegede, 2001).

In contrast, gender was not a significant factor in predicting the learners’ self-confidence in using computers (Contreras, 2004), and the use of the computers and the Internet (Atan, Sulaiman, Rahman, & Idrus, 2004). Other studies have reported that previous computer experiences and numbers of online courses taken better predicted the students’ learning outcomes (Lim & Kim, 2003) and confidence in using computers (Contreras, 2004) in an online learning environment.

In reviewing the existing research involving learner variables, we noted several issues that provide connections between previous research and this current study. First, while a number of the studies provided meaningful findings, a limitation of these studies was that the learning environments only included the usage of a single delivery method (e.g., online instruction, computer-based instruction, or classroom instruction) not multiple or blended delivery methods. Few studies have been conducted to examine the influence of learner variables on students’ learning within a blended learning environment. Second, there were few existing studies that utilized integrated approaches to
examine the empirical assessment of the influence of diverse learner variables on students’ learning within a blended learning environment. From this review, there is a pressing interest among researchers to identify the mediating mechanisms of contextual characteristics of learner variables influencing students’ learning and learning application within in a blended learning environment. This study will address these issues.

**Instructional Variables**

A thorough review of the studies empirically examining instructional variables has suggested that at least four categories of instructional design constructs exist: (a) quality of instructor, (b) quality of learning activity, (c) learning support, and (d) study workload. In terms of instructor quality, instructor’s command of subject matter and ability to make clear and understandable presentations during class were frequently cited among studies (Lim, 2002). Researchers also have examined several salient instructor variables such as the use of more applicable learning examples during lectures (Axtell, Maitlis, & Yearta, 1997), inclusion of general rules and principles for learners to apply to the learned content (Goldstein, 1986), and a greater specificity of instructional content to be applied in other settings (Clark & Voogel, 1985).

Concerning improvement in the quality of learning activities, research studies have suggested several instructional strategies that include: (a) planned goal-setting prior to the learning experience (Wexley & Nemeroff, 1975), (b) utilization of action planning methods (Foxon, 1997), and (c) the inclusion of various instructional methods like application examples in various contexts, the use of analogies, and the usage of computer simulations during the learning process (Garavaglia, 1993). As an indicator of quality of learning activities, research involving students’ satisfaction with instructional programs has indicated seven instructional areas influence satisfaction. Those areas include: satisfaction with instructor, instructional methods/activities, learning objectives, logistical matters, topics/content, planned action/transfer, and course materials (Sanderson, 1995). Finally, study workload, has been empirically examined as an instructional design construct in a few studies. In comparing students’ study time between online versus classroom instruction Oh and Lim (2005) found that online students spent more time completing one week’s course workload than those with classroom-based courses.

**Learning Motivation**

Within learning research, learning motivation is one of the most frequently studied variables. The concept of learning motivation has been defined as the organized pattern of pursuing goals, beliefs, and emotions (Ford, 1992). Wlodkowski (1985) defined learning motivation as a force to arouse, give direction to, continue, and choose a particular learning behavior. In examining the relationship of learning motivation on classroom-based instruction and other types of instruction such as computer-based and online instruction, Sankaran and Bui (2001) found that less motivated learners did not perform as well on knowledge tests as motivated students. Students with high learning motivation tended to engage in more meta-cognitive strategies and were more likely to persist at a task than students with low learning motivation (Pintrich & De Groot, 1990). Similarly, Salili, Chiu, and Lai (2001) found that students who were confident and motivated to learn, spent more time and effort and achieved higher levels of performance than those who were not confident and motivated.

In their assessment of online learning motivation, Lim and Kim (2003) created a typology of six learning motivation variables that included: (a) reinforcement, (b) course relevance, (c) interest, (d) self-efficacy, (e) affect, and (f) learner control. Regarding reinforcement, several theorists have noted reinforcement is important to learning motivation (Graham, 1994; Weiner, 1994). Relevance refers to the value of course content related to learner’s jobs and studies. Atkinson (1964), Herzberg, Mausner, and Snyderman (1959) indicated the perceived level of value residing in learning content decides the level of motivation. According to Lim and Kim (2003), level of interest is another type of motivation factor promoting learner involvement during learning. When a learning task is challenging and involves fantasy during the learning process, learners will be motivated (Malone, 1981). The fourth type of learning motivation is self-competence which refers to one’s feelings of self-worth and self-efficacy. Self-competence can be generally described as the degree to which one believes that he or she is able to achieve a given task (Foster, 2001). The fifth type of motivation, affect, involves a learner’s feeling and emotion during the learning experience (Wlodkowski, 1999). Lastly, Lim and Kim suggested learner control as the sixth type of motivation. Learner control has been an important study subject in the field of instructional systems and technology (Chung & Davis, 1995; Reigeluth, 1989). As noted earlier, although the Lim and Kim (2003) typology was originally
developed to assess online learning motivation, the typology might have application for examining blended learning environments. This study will examine this possibility.

Learning Involvement

Current research also suggests that learning involvement is another variable that influences students’ learning and satisfaction within traditional and/or online learning environments. In this study, learning involvement is defined as the degree to which learners interact with other learning component (i.e., learning content, learning activities, peer learners, tutors, and instructors) and are engaged in the learning process. In terms of learning involvement, Sankaran and Bui (2001) found that students who used undirected strategies demonstrated lower levels of performance than those who used ‘deep’ or ‘surface’ strategies, such as taking notes or practicing exercises during online learning. Study groups and contacts with the instructor were also reported as positively contributing to student success (King, Harner, & Brown, 2000). Other researchers have emphasized the importance of online learners’ psychological state in blended and online learning environments, especially since learners may feel unsupported and experience an increased workload if they lack the sense of presence or belonging during learning (McMillan & Chavis, 1986).

Clearly, learning involvement is a variable of interest to be explored in examining learning outcomes associated with a blended instruction delivery format.

Purpose and Construct of the Study

The purpose of this study is to examine how student’s learning in a blended learning environment is influenced by learner’s demographic characteristics and learning styles, instructional design, and learning motivation and involvement. Rationales for inclusion of each of the three antecedents as predictors of student learning in a blended learning environment are now given.

In this study, instructional design was operationally defined into four dimensions: (1) the quality of instructor, (2) learning activities, (3) learning support, and (4) study workload. Learning motivation and involvement comprised another set of antecedent constructs of our study as previous research studies supported the distinctive nature of these variables in influencing students’ learning and learning application (Ford, 1992; Wlodkowski, 1985).

Regarding the dependent variables of our study, we purposefully selected multiple categories of learning outcomes: students’ actual learning, student’s perceived learning, and perceived learning application. Table 1 presents the study variables we included in this study.

<table>
<thead>
<tr>
<th>Table 1. Operational variables of the study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent variables</strong></td>
</tr>
<tr>
<td>Learner variables</td>
</tr>
<tr>
<td>- Gender</td>
</tr>
<tr>
<td>- Age</td>
</tr>
<tr>
<td>- Distance learning experience</td>
</tr>
<tr>
<td>- Preference in delivery format</td>
</tr>
<tr>
<td>- Average study time</td>
</tr>
<tr>
<td><strong>Instructional variables</strong></td>
</tr>
<tr>
<td>- Instructor quality</td>
</tr>
<tr>
<td>- Learning activities</td>
</tr>
<tr>
<td>- Learning support</td>
</tr>
</tbody>
</table>
| - Study workload                            | **Table 1. Operational variables of the study**

Methodology

Samples

In order to assess student learning outcomes involving a blended instruction learning environment, a group of undergraduate students enrolled in a learner and program evaluation course at a southeastern university were asked to participate in this study. The subjects for the study included 60 students (21 male and 39 female). Among the 60 students, 38 were freshmen or sophomore students while 22 were junior or senior students.
Instrument and Procedure

Dependent Variables: Student Learning Outcomes

Instrumentation assessing three forms of learning outcomes (actual learning, perceived learning, and perceived learning application) was used in this study to measure student learning outcomes for eighteen course-related learning objectives. Actual learning was assessed utilizing an in-class examination including a set of content items designed to assess actual learning gain before and after each semester. In order to assess perceptions of learning and perceive application of learning, a questionnaire using a five point Likert-type scale to measure the perceived degree of learning (1 "do not understand" to 5 "completely understand") and the perceived application of learning (1 "none" to 5 "frequently use") was created to assess changes for the eighteen learning objectives of the course used throughout the semester. To assess students’ perceived increase in learning, this study collected both pre- and post-perception data about students’ learning before and immediately after the course at each semester.

The open-ended part of the questionnaire asked the students’ responses about the reasons for high or low perceived learning and learning application. Other questions within the open-ended part of the questionnaire included questions asking the students’ comments and suggestions to improve the course. Questions related to students’ involvement with the learning activities and peer students were also included in the open-ended part of the questionnaire.

Independent Variables: Instructional Design, Learning Involvement, and Motivation

Instrumentation using 17 closed-ended questions assessing learners’ satisfaction with instructional factors such as quality of instructor, learning activities, and learning support, and the learners’ perception of study workload was used. All question items in this category used a five point Likert-type scale. Regarding learning involvement, eight question items asking students’ perceived involvement in the areas of interest in subject content, learning progress, learning involvement, personal effort, preparedness, and personal challenge were included.

Prior to conducting the data analysis, factor analysis was performed to screen variables within the student learning outcomes, instruction design, and learning motivation and involvement factors that may have causal relationships in predicting learning outcomes. Results from the factor analysis indicated that all scales within each variable fell into one single factor. Overall, Cronbach’s alphas were .93 for the perceived learning, .94 for the learning application, .70 for the test, .93 for the instructor quality, .81 for the learning activities, .85 for the learning support, .90 for the workload, and .90 for the learning involvement respectively. To collect the pre- and post-survey data, the students were asked to participate in the surveys conducted online at the beginning and at the end of each semester. Data collection was performed for four semesters between 2004 and 2005.

In order to assess learning motivation, the study utilized the Learning Motivation Questionnaire (LMQ) composed of 24 question items representing the six sub categories (course relevancy, course interest, affect/emotion, reinforcement, self-efficacy, and learner control) (Lim & Kim, 2003). The LMQ used a five point Likert-type scale to measure the level of learning motivation (1 “strongly disagree” to 5 “strongly agree”) and the data collection was completed via online at the end of each semester. Results from the factor analysis also indicated that all sub scales within the LMQ fell into one single factor for the six motivation types respectively. The reliability alpha was .92 for the LMQ.

Data Analysis

For data analysis, we have utilized both quantitative and qualitative methods to analyze various data sets acquired for this study.

Quantitative Analysis

Basic descriptive statistics were used to analyze the test scores and the perceived degree of learning, application of learning, and instructional quality of the course. ANOVA was used to assess the differences in perceived learning, posttest, and learning application means between the comparison groups based on learner variables. Repeated measures ANOVA was used to test if differences exist for the perceived learning and actual learning increase over the time period between before and after each semester based on the learner variables. Correlation analyses were
performed to test inter-relationships among those variables in instructional variables and learning outcomes. Finally, multiple regression analysis was conducted to assess the influence of learner, instructional, and motivational variables on the students’ learning outcomes. Prior to the regression analysis, factor analysis was used to screen variables within the instructional factors and motivational variables that may have causal relationships.

Qualitative Analysis

The investigators in this study conducted domain analyses employing content analysis procedures for qualitative data (Spradley, 1979). This method sorted through the open-ended responses, and identified themes and patterns in the reasons that promote or hinder the learners’ learning and application during learning. After content analysis, cumulative frequencies and percentages for similar types and attributes identified in the domain categories were calculated to determine how often similar types were elicited. This allowed the investigators to determine which terms were elicited most frequently and to gain a better understanding about the distribution of beliefs across domain categories.

Context of the Course

The course used in this study was developed to teach curriculum content in learner and program evaluation to undergraduate students. The course was delivered through a blended delivery method for two years utilizing both classroom instruction and online delivery methods. In order to best address the blended nature of the instruction, the instructor adopted a pedagogical framework involving four major components: (1) self-paced and individualized learning through online modules to understand the basic subject content of the course, (2) instructor-led instruction to have more in-depth understanding of the learned content, (3) group collaboration sessions to widen their views and application opportunities of their learning, and lastly (4) checking for students’ learning and application through weekly and semester-based assessment of their learning progress.

Regarding the blending mix of each delivery format, about half of the instruction was conducted in class and the other half was delivered through online delivery methods. For the effective delivery of the blended instruction, the students were required to attend weekly classroom instruction that the instructor presented each week’s major learning content and utilized various learning activities for learning involvement and collaboration (e.g., demonstration, case studies, group discussions, etc.). After attending each week’s classroom instruction, the students were asked to complete online learning modules to reinforce their classroom learning.

Findings

Differences in Learning Outcomes, Instructional Quality, and Learning Motivation

For the purpose of this study, the researchers categorized learner variables into gender, age, distance learning experience, preference in delivery format, and study time. Data analysis indicated that the learners with different learner variables seemed to have different learning outcomes and perceptions about the instructional quality and learning motivation and involvement. Tables 2 shows the mean scores of learning outcome variables (i.e., perceived learning, perceived learning increase, posttest, actual learning increase, perceived learning application) based on learner variables.

In Table 2, ANOVA results are shown regarding what learning outcomes were significantly different based on each learner variable. From the ANOVA results, gender was not indicated as a learner variable differentiating any of the sub variables of the students’ learning outcomes in this study. For age, learners between 20 years or higher (junior or senior students) had significantly higher mean scores in perceived learning, learning application, and learning involvement than other age groups (freshman and sophomore students). Learners with distance learning experience indicated significantly lower mean scores for the quality of learning activity and learning motivation than those who didn’t have distance learning experience. Regarding preference in delivery format, those learners who preferred online learning method showed significantly higher mean scores for perceived learning, learning application, learning activity, learning motivation, and learning involvement than those who did not. For study time, a post-hoc
test revealed that those students who studied up to one hour indicated a significantly higher mean score for general workload than those who studied more than 2 hours (p=.038), while those students who studied up to one hour indicated a significant higher mean score for learning involvement than those who studied between one and two hours (p=.037).

Table 2. Mean scores and ANOVA results according to learner variables

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub category</th>
<th>N</th>
<th>Perceived Learning P</th>
<th>Perceived Learning Increase P</th>
<th>Postest p</th>
<th>Actual Learning Increase p</th>
<th>Learning Application p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>20</td>
<td>3.98/.54</td>
<td>.791</td>
<td>.81/.63</td>
<td>10.84/3.63</td>
<td>3.32/3.70</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>39</td>
<td>3.73/.51</td>
<td>.52/.69</td>
<td>.52/.69</td>
<td>11.25/3.72</td>
<td>2.48/3.29</td>
</tr>
<tr>
<td>Age</td>
<td>18-19</td>
<td>38</td>
<td>3.69/.53</td>
<td>.46/.64</td>
<td>.49/.64</td>
<td>11.03/3.53</td>
<td>2.39/3.11</td>
</tr>
<tr>
<td></td>
<td>20 and under</td>
<td>21</td>
<td>4.03/.46</td>
<td>.85/.70</td>
<td>.85/.70</td>
<td>11.29/3.98</td>
<td>3.38/3.92</td>
</tr>
<tr>
<td>Distance Learning Experience</td>
<td>Yes</td>
<td>23</td>
<td>3.76/.64</td>
<td>.408</td>
<td>.65/.67</td>
<td>10.17/3.51</td>
<td>3.61/2.94</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>37</td>
<td>3.85/.45</td>
<td>.60/.70</td>
<td>.57/.70</td>
<td>11.72/3.68</td>
<td>3.47/.54</td>
</tr>
<tr>
<td>Preference in delivery format</td>
<td>Online</td>
<td>33</td>
<td>4.02/.44</td>
<td>.001</td>
<td>.67/.61</td>
<td>11.88/3.88</td>
<td>3.22/3.42</td>
</tr>
<tr>
<td></td>
<td>F2F</td>
<td>26</td>
<td>3.55/.52</td>
<td>.56/.76</td>
<td>.56/.76</td>
<td>10.15/3.27</td>
<td>3.00/3.31</td>
</tr>
<tr>
<td>Average Study Time</td>
<td>&gt; 1 hr</td>
<td>27</td>
<td>3.69/.56</td>
<td>.64/.62</td>
<td>.64/.62</td>
<td>11.33/3.36</td>
<td>2.63/2.26</td>
</tr>
<tr>
<td></td>
<td>&lt; 2 hrs</td>
<td>8</td>
<td>3.78/.49</td>
<td>.77/.96</td>
<td>.77/.96</td>
<td>9.63/3.78</td>
<td>2.63/5.40</td>
</tr>
</tbody>
</table>

Table 3. Learner variables over the different time periods

<table>
<thead>
<tr>
<th>Category</th>
<th>Interaction B/W Time</th>
<th>Main Effect</th>
<th>Interaction B/T</th>
<th>Main Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>.121</td>
<td>.408</td>
<td>.382</td>
<td>.251</td>
</tr>
<tr>
<td>Age</td>
<td>.050*</td>
<td>.184</td>
<td>.292</td>
<td>.741</td>
</tr>
<tr>
<td>Distance Learning Experience</td>
<td>.798</td>
<td>.368</td>
<td>.040*</td>
<td>.373</td>
</tr>
<tr>
<td>Preference in delivery format</td>
<td>.541</td>
<td>.001</td>
<td>.177</td>
<td>.104</td>
</tr>
</tbody>
</table>

*Repeated measures ANOVA
*p<.05 (Two-tailed tests)
When comparing the students’ mean scores of the perceived learning increase and pre/posttest between before and after each semester, the repeated measures ANOVA revealed several meaningful findings about the differences among the learner variables. As Table 3 presents, students who were 20 years old or more (i.e., junior and senior students) indicated significant differences on overall perceived learning increase scores before and after each semester than those who were between 18 and 19 years old. Regarding the increase in mean scores between pre- and post-test, students without prior distance learning experience indicated significantly higher mean increase scores than their student counterparts.

**Influence of Instructional and Learner Variables on Learning Outcomes**

In analyzing the relationships between instructional and learner variables and the students’ learning outcomes, several meaningful findings were obtained. First, the students’ mean score of perceived learning was significantly related with the mean scores in instructor quality, learning activity, learning support, learning motivation, and learning involvement. Similarly, the students’ perceived learning application mean score was significantly related with the mean scores in instructor quality, learning activity, learning support, learning motivation, and learning involvement. Other mean scores of learning outcomes (i.e., perceived learning increase, actual learning, actual learning increase) were not significantly related with any of the study variables, with the exception of one relationship between perceived learning increase and learning activity. Table 4 reports the analysis results.

**Table 4. Correlations between learning outcomes and instructional and learner factors**

<table>
<thead>
<tr>
<th></th>
<th>Instructor quality</th>
<th>Learning activity</th>
<th>Study workload</th>
<th>Learning support</th>
<th>Learning motivation</th>
<th>Learning involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived learning</td>
<td>.477**</td>
<td>.618**</td>
<td>.105</td>
<td>.319</td>
<td>.506**</td>
<td>.456**</td>
</tr>
<tr>
<td>Perceived learning increase</td>
<td>.203</td>
<td>.395*</td>
<td>.060</td>
<td>.149</td>
<td>.185</td>
<td>.175</td>
</tr>
<tr>
<td>Actual learning</td>
<td>-.059</td>
<td>.072</td>
<td>-.074</td>
<td>-.171</td>
<td>.099</td>
<td>-.050</td>
</tr>
<tr>
<td>Actual learning increase</td>
<td>-.022</td>
<td>.075</td>
<td>-.025</td>
<td>-.116</td>
<td>.262</td>
<td>.066</td>
</tr>
<tr>
<td>Perceived learning application</td>
<td>.528**</td>
<td>.563**</td>
<td>.226</td>
<td>.339**</td>
<td>.386**</td>
<td>.493**</td>
</tr>
</tbody>
</table>

*p<.05,  **p<.01 (Two-tailed tests)*

Among the six types of learning motivation, some types appealed to influence the students’ learning more than others. As presented in Table 6, all motivation types were significantly correlated with the students’ perceived learning. However, for the relationship between perceived learning increase and post-test, no significant relationships with any of the motivation types were found. Course relevancy was the only motivation type that was significantly related with actual learning increase. Among the six learning motivation types, affect, reinforcement, self-efficacy, and learner control were significantly correlated with the students’ perceived learning application. As expected, learning involvement was significantly related with all motivation types. The students’ average study time was significantly related with relevancy, interest, and affect.

**Table 5. Correlations between learning outcomes and motivation types**

<table>
<thead>
<tr>
<th></th>
<th>Relevancy</th>
<th>Interest</th>
<th>Affect</th>
<th>Reinforcement</th>
<th>Self-efficacy</th>
<th>Learner control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived learning</td>
<td>.354**</td>
<td>.384**</td>
<td>.456**</td>
<td>.342</td>
<td>.518**</td>
<td>.430**</td>
</tr>
<tr>
<td>Perceived learning increase</td>
<td>.177</td>
<td>.200</td>
<td>.190</td>
<td>.050</td>
<td>.059</td>
<td>.195</td>
</tr>
<tr>
<td>Posttest</td>
<td>.224</td>
<td>-.014</td>
<td>.043</td>
<td>-.042</td>
<td>.145</td>
<td>.178</td>
</tr>
<tr>
<td>Actual learning increase</td>
<td>.367**</td>
<td>.171</td>
<td>.215</td>
<td>.107</td>
<td>.206</td>
<td>.230</td>
</tr>
<tr>
<td>Learning application</td>
<td>.264</td>
<td>.265</td>
<td>.340*</td>
<td>.354**</td>
<td>.396**</td>
<td>.282*</td>
</tr>
<tr>
<td>Learning involvement</td>
<td>.584**</td>
<td>.573**</td>
<td>.583**</td>
<td>.399**</td>
<td>.415**</td>
<td>.519**</td>
</tr>
<tr>
<td>Average study time</td>
<td>.267*</td>
<td>.350**</td>
<td>.339*</td>
<td>.071</td>
<td>.127</td>
<td>.060</td>
</tr>
</tbody>
</table>

*p<.05,  **p<.01 (Two-tailed tests)*
While the method of correlation analysis has been widely used to detect specific relationships between study variables, researchers have previously recommended using multiple regression analyses to establish the relative predictive importance of the independent variables (Allison, 1999). When we conducted the multiple regression, the findings about the influence of learner, instructional, and motivational variables on learning outcomes were consolidated around one variable in perceived learning application \( R^2 = .548, p = .001 \), while all other variables indicated insignificant \( R^2 \) values (.195 for the perceived learning, .269 for the perceived learning increase, .228 for the post-test, and .240 for the actual learning increase). In order to detect which variables influence students’ perceived learning application, stepwise regression analysis was performed. Table 6 includes the results of the stepwise regression.

Table 6. Stepwise regression for perceived learning application

<table>
<thead>
<tr>
<th>Learning outcomes</th>
<th>( R^2 )</th>
<th>Predictor</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived learning application</td>
<td>.469</td>
<td>Age</td>
<td>.291*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of instructor</td>
<td>.217*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learning activity</td>
<td>.218*</td>
</tr>
</tbody>
</table>

* \( p < .05 \)

Reasons for High or Low Learning and Learning Application

From the students’ responses to the open-ended questions regarding the reasons to facilitate or inhibit perceived learning and learning application, various themes were observed. In the online survey, the students were asked to provide the three most influential reasons supporting or hindering their perceived learning and learning application. Collectively, the reasons associated with instructional effectiveness were identified as the most influential factor for students’ learning (50%). Some examples under the instructional effectiveness category included usefulness of class assignment and projects, clear and concise online learning content, review and repetition of learning, and useful learning activities. Table 7 reports verbatim reason categories given for the students’ high and low perceived learning along with the frequencies and percentages of the reason categories.

Table 7. Reasons for high or low learning and learning application

<table>
<thead>
<tr>
<th>Reason Category</th>
<th>( f )</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reasons for high learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional effectiveness</td>
<td>21</td>
<td>18.75</td>
</tr>
<tr>
<td>Previous learning</td>
<td>14</td>
<td>12.50</td>
</tr>
<tr>
<td>Related to my current or future jobs</td>
<td>7</td>
<td>6.25</td>
</tr>
<tr>
<td>High interests in the learning content</td>
<td>6</td>
<td>5.36</td>
</tr>
<tr>
<td>Personal learning effectiveness</td>
<td>4</td>
<td>3.57</td>
</tr>
<tr>
<td>Opportunity to practice learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal motivation for learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>112</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><strong>Reasons for low learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional ineffectiveness</td>
<td>20</td>
<td>43.48</td>
</tr>
<tr>
<td>Lack of interest in the learning content</td>
<td>9</td>
<td>19.57</td>
</tr>
<tr>
<td>Lack of personal effort</td>
<td>3</td>
<td>6.52</td>
</tr>
<tr>
<td>Lack of understanding</td>
<td>3</td>
<td>6.52</td>
</tr>
<tr>
<td>Not related to my work</td>
<td>3</td>
<td>6.52</td>
</tr>
<tr>
<td>Lack of opportunity to use learning</td>
<td>2</td>
<td>4.35</td>
</tr>
<tr>
<td>Interruptions during learning</td>
<td>1</td>
<td>2.17</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>46</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><strong>Reasons for high learning application</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opportunity to use learning</td>
<td>38</td>
<td>40.86</td>
</tr>
<tr>
<td>Applicable learning content to my work or career</td>
<td>19</td>
<td>20.43</td>
</tr>
<tr>
<td>Because of high learning</td>
<td>18</td>
<td>19.35</td>
</tr>
<tr>
<td>Experience from previous learning</td>
<td>7</td>
<td>7.53</td>
</tr>
</tbody>
</table>

290
<table>
<thead>
<tr>
<th>Reason for low learning application</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of understanding of learning content</td>
<td>11</td>
<td>28.95</td>
</tr>
<tr>
<td>Not related or applicable to my job</td>
<td>9</td>
<td>23.68</td>
</tr>
<tr>
<td>Not enough opportunity to use during class</td>
<td>6</td>
<td>15.79</td>
</tr>
<tr>
<td>Lack of opportunity to use learning in my job</td>
<td>4</td>
<td>10.53</td>
</tr>
<tr>
<td>Lack of clear instruction for application</td>
<td>2</td>
<td>5.26</td>
</tr>
<tr>
<td>Too much content to apply for a given time</td>
<td>2</td>
<td>5.26</td>
</tr>
<tr>
<td>Not stressed to apply</td>
<td>2</td>
<td>5.26</td>
</tr>
<tr>
<td>Learning activities were not related</td>
<td>1</td>
<td>2.63</td>
</tr>
<tr>
<td>Lack of motivation to apply</td>
<td>1</td>
<td>2.63</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>38</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Conclusions and Implications

A review of the existing literature indicates there is considerable interest among researchers and practitioners for identifying when and how blended instruction is working or is not working, especially in the delivery of college level courses (Arbaugh, 2005). From the findings of this study, two major themes to improve the current practices of blended instruction in college education seemed to emerge. First, individual learning differences are an important area to consider in making instructional decisions for learner-oriented blended instruction. Researchers have claimed that most theories of instructional design fail because much of the evolving research on individual learning differences remains broadly focused on cognitive interests and intrinsic mechanisms for information processing and knowledge building (Martinez & Bunderson, 2000). To address the issue of individual differences in learning, two levels of instructional decisions should be involved. The existing literature suggests making ‘macro’ and ‘micro’ level instructional decisions was considered a critical task in designing any type of instruction associated with traditional and blended learning environments (Hannafin, Land, & Oliver, 1999; Hirumi, 2002). An example of the macro level instructional decision is the identification of instructional strategies for blended instruction and accommodating individual learning differences to decide blending mix among different delivery methods. An example of the micro level decisions are examining the types and roles of interactions and learning activities and assessing the effect of learner characteristics such as demographic information, learning styles/preferences, and learners’ learning motivation prior to attending blended instruction courses.

The second theme evolving from this study’s findings relates to the level of student learning. In college education, promoting a higher level of learning has become a prime interest among instructors as more universities and colleges try to develop and deliver quality blended instructional courses. Making students’ learning experiences applicable, in this sense, is an important task that today’s instructors in college education should address from design stage to delivery and evaluation stages of blended instruction. From the study findings applicability of learning content seems to be one critical factor of instruction design to sustain students’ learning interest and promote learning increase during the blended instruction. From the researchers’ personal experiences, designing an instruction with a focus on learning application seemed to require different approaches in instructional needs assessment and selecting learning activities for instructional design than those focusing on learning level only. When focusing on learning, instructional designers, as an example, may use learning activities that aid learning of facts, concepts, procedures, and skills based on certain performance criteria. When focusing on learning application, instructional designers may want to identify and target students’ learning application needs first, and then, identify appropriate learning activities that allow more opportunities for learning transfer. To create meaningful learning experience allowing applicable moments of student learning, several instructional strategies are suggested: matching instructional level to students’ competency level, use of identical elements between learning and application settings (Baldwin & Ford, 1988), and variation in the modality of instructional delivery to promote near and far transfer of learning (Clark & Taylor, 199).
Limitations and Future Studies

Several limitations in this study exist. First, this study utilized students’ perception data to assess the effect of learner and instructional variables on students’ learning. Second, the selection and sample frame used in the study limits the generalization of the study findings. Although we collected the data from groups of students who took a college course between 2004 and 2005, the size of the sample somewhat limits the generalizability of the findings. The data from a single population in a university setting may also reflect the institutional characteristics to the study findings. The subjects’ exposure to the course content in program evaluation may produce some sensitizing effect for them to respond to surveys conducted after completion of the course so it may limit the generalization of the study findings.

The primary research objective of this study was to identify the effect of learner and instructional variables on students’ learning and perceived learning application in a blended instruction learning environment. We used a variety of data sets to examine learning differences among students’ groups based on the learner and instructional variables and to determine the influence of such variables on learning outcomes. Verification of our study results from other learner and instructional variables (e.g., learning styles, technical skills), however, might be worthy of important consideration in future research studies. In terms of generalization, future studies will need to use a broader target population. Future studies may reveal similar findings, particularly if replication procedures are conducted in a different blended instruction setting (e.g., different blending mix of instructional methods and time).

References


Technology's Role in Field Experiences for Preservice Teacher Training

Emily Hixon1 and Hyo-Jeong So2

1School of Education, Purdue University Calumet, 2200 169th St., Hammond, IN 46323, USA // Tel: +1 219 989-2150 // Fax: +1 219 989-3215 // hixone@calumet.purdue.edu
2National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616 // Tel: +65 6790-3283 // Fax: +65 6896-8038 // hyojeong.so@nie.edu.sg

ABSTRACT

The purpose of this paper is to provide a comprehensive review of how technology has been used to enhance or replace field experiences in preservice teacher preparation programs, and discuss the benefits and limitations of traditional and technology-enhanced/virtual field experience approaches. In this paper, three types of technology-enhanced field experiences are discussed: Type I - concrete, direct experience in reality; Type II – vicarious, indirect experience with reality; and Type III – abstract, experience with model of reality. Five specific benefits of technology use in field experiences are identified, namely a) exposure to various teaching/learning environments, b) creation of shared experiences, c) promoting reflectivity, d) preparing students cognitively, and e) learning about technology integration. Several limitations of technology-integrated field experiences are also discussed, including (a) lack of interaction with teachers and students, (b) limited reality and complexity, (c) availability of relevant cases, and (d) technical problems. In conclusion, it is suggested that the overall goals and objectives for a specific field experience must be the focus when field experience options are being explored.

Keywords

Field experiences, Educational technology, Teacher education programs, Preservice teacher education, Distance education

Introduction

Field experiences are a foundation in preservice teacher education programs today. In fact, early teacher educators have found value in the “learning by doing” approach since as early as the mid-nineteenth century (Cruickshank & Armaline, 1986). Consistent with John Dewey’s emphasis on experiential education, field experiences today are focused on providing examples of best practices and pairing students with teachers who are not only excellent teachers, but also excellent role models willing to engage in reflective practice with preservice teachers (Frieberg, 1995; Posner, 2005).

Numerous benefits of field experiences in teacher preparation programs have been documented in the literature. They include helping students decide if teaching is the right career choice, providing an opportunity for students to practice skills prior to student teaching, helping preservice teachers start viewing themselves as teachers, and improving preservice teachers' attitudes toward teaching (McIntyre, 1983). Field experiences are also typically offered in conjunction with a course as a way of helping students better understand the conceptual and theoretical knowledge being presented (Hopkins, 1995). Recently, much attention has also been paid to the role that field experiences may play in helping preservice teachers learn how to effectively integrate technology into their teaching (e.g., Bahr, Shaha, Farnsworth, Lewis, & Benson, 2004; Dawson & Dana, 2007; Evans, 2004; Wentworth, Graham, & Tripp, 2008).

These outcomes of field experiences are clearly desirable and emphasize the importance of field experiences in teacher education. While field experiences are certainly a valuable component of teacher education programs, factors surrounding their implementation may influence the quality and impact of field-based experiences. Of major concern to teacher educators is the fact that prospective teachers may not be cognitively prepared to benefit from experiences in the field, especially when they occur early in their education program (Feiman-Nemser & Buchmann, 1985; Goodman, 1986; Hannah, 1995). When student teachers are not cognitively prepared to learn from experiences in the field, they sometimes view the field experience as an off-campus activity as opposed to on-the-job training, and believe that the field experiences do not provide “real teaching experiences” (Aiken & Day, 1999).

A similar concern is that students are not taking an active role in developing, processing, and reflecting on their field experiences (Cruickshank & Armaline, 1986; Feiman-Nemser & Buchmann, 1985; Goodman, 1986; Johnston, 1995).
Although reflective thinking has been emphasized in teacher education programs (e.g., Brookfield, 1995), it has been documented that students often participate in field experiences with a very mechanical, “learn the skill” approach and are not thoughtful about the experience as it relates to their education. In addition to promoting reflective inquiry, education faculty may find it difficult to facilitate appropriate reflection and debriefing related to field experiences because students do not share a common experience to serve as a jumping-off point for discussions (Hannah & Abate, 1992–1993). The varied experiences of students in the field also make it difficult to ensure consistent quality of student field experiences.

On a more practical note, with the impending teacher shortage and growing preservice education programs in some countries, especially the United States, schools of education are being forced to place students at sites increasingly distant from campus. This puts financial and time burdens on both the school (who must appropriately supervise) and the students (e.g., Garrett & Dudt, 1998; Knight, Pedersen, & Peters, 2004). A related concern, especially for colleges and universities located in rural and less diverse areas, is how to ensure students gain experience with diverse populations (e.g., Lehman & Richardson, 2003).

In an effort to address some of the afore-mentioned concerns yet preserve the benefits associated with traditional field experiences, educators are exploring the role that various technologies may play in supplementing and perhaps replacing traditional early field experiences. As schools consider the inclusion of technology-enhanced and/or virtual field experiences in their program, they need to understand the strengths and limitations of such approaches, and consider their viability based on the purpose of the field experience. The purpose of this paper, therefore, is to provide a comprehensive review of how technology has been used to enhance and sometimes replace field experiences in preservice teacher education programs, and discuss the benefits and limitations of both traditional and technology-enhanced/virtual field experience approaches.

To examine technology’s role in field experiences, relevant studies in the literature of teacher education were analyzed. Library databases (e.g., Academic Search Premier and ERIC) and Google Scholar were used to conduct a search using various combinations of the following keywords: “field experiences”, “early field experiences”, “practicum”, “technology”, “virtual”, “video” and “preservice teacher”. Since the focus of the present paper is on the role of technology in field experiences, studies that (a) discuss general issues of technology integration in preservice teacher education, or (b) do not have a connection to a field experiences component, were excluded in the search process. The constant comparative method (Glaser & Strauss, 1967; Lincoln & Guba, 1985) was used to identify and analyze emerging themes across studies. These common themes were then categorized in terms of types, benefits, and limitations, which are discussed in detail in this report.

Technology–Enhanced and Virtual Field Experiences

There are a variety of ways that technology can be implemented in conjunction with field experiences. Presented here are three types of technology-enhanced field experiences, which is an adaptation of the classification scheme put forth by Paese (1996). As shown in Figure 1, the three types can be categorized according to the degree to which experiences are situated in reality.

![Figure 1. Three types of technology-enhanced field experiences](image-url)
Type I experiences involve preservice teachers being placed in real classrooms where they teach and/or observe real students as in the traditional field-based approach. Technology tools are used to facilitate supervision, reflection, and/or communication. Numerous studies have examined such uses of technology in conjunction with field-based experiences (e.g., Bonk, Malikowski, Angeli, & East, 1998; Doering, Johnson, & Dexter, 2003; Edens, 2000; Garrett & Dudit, 1998; Gruenhagen, McCracken, & True, 1999; Holstrom, Ruiz, & Weller, 2007; Johnson, Maring, Doty, & Fickle, 2006; Kale, Hur, Yerasimou, & Brush, 2006; Knowlton, 2004; Laffey & Musser, 1998; Liu, 2005; Mason, 2000; Moffett, 2001; Whipp, 2003).

In Type II experiences, preservice teachers gain vicarious experience by remotely observing teachers and students in real classrooms. Two technical options are available in this category. Video-conferencing technology supports synchronous observations of classroom lessons and possibly even interaction with the teacher and/or students, while pre-recorded video cases available on CD-ROMs or via the Internet allow for non-real-time observations. Type II experiences have been described and investigated by several researchers (e.g. Atkins, 1998; Baker, 2005; Green, 2005; Hannah, 1995; Hannah & Abate, 1992–1993; Knight et al., 2004; Lambdin, Duffy, & Moore, 1996; Lehman & Richardson, 2003; Malewski, Phillion, & Lehman, 2005; McDevitt, 1996; McIntyre & Pape, 1993; Putnam & Borko, 2000; Santagata, Zannoni, & Stigler, 2007).

Type III field experiences utilize simulated environments. Zibit and Gibson (2005) call this type of field experience as a “virtual practicum” based on simulated apprenticeship models. Instead of teaching or observing real teachers and students, preservice teachers experience an artificial model of reality by observing and/or teaching virtual teachers and students in a simulated classroom. Unlike the technologies used in Types I and II experiences, there is very limited research on the use of simulated classrooms as field experience components. Technical difficulties and cost issues for developing simulated environments and learners with artificial intelligence functions seem to be major obstacles explaining the lack of research in this area (Brown, 1999; Payr, 2005). Among the few studies on simulated classroom environments, Foley and McAllister (2005), Ferry and colleagues (2005), and Girod and Girod (2006) have explored possibilities of linking simulations and field experiences.

While the three types of field experiences are discussed separately above, it is also common for different types to be used in conjunction with one another. Preservice teachers who remotely observe a classroom using technology (Type II experience) or interact with a simulated classroom environment (Type III experience) may also be participating in a field-based experience that may or may not utilize some form of technology for communication, supervision, etc. (Type I technology-enhanced field experience or a strictly traditional field experience). This blended approach seems to help students cognitively prepare for real teaching experiences, and transfer knowledge and skills learned in virtual environments to real classrooms (e.g., Baker, 2005; Lee & Powell, 2005–2006a; Lee & Powell, 2005–2006b; Liu, 2005).

**Benefits of Technology Use**

Providing preservice teachers with opportunities to observe and interact with classroom environments, real or simulated, is critically important to educating high quality, well-prepared teachers. While technology's role in relation to field experiences is still being explored, it is important to identify the potential benefits of technology's use that have already been documented in the literature. Such information is relevant to teacher educators and others involved in preservice teacher education to ensure that they design the most effective field experiences given their goals, situations, and resources. Five benefits of technology use in field experiences are presented in this section: a) exposure to various teaching/learning environments, b) creation of shared experiences, c) promoting reflectivity, d) preparing students cognitively, and e) learning about technology integration.

**Exposure to Various Teaching and Learning Environments**

Students participating in a traditional field experience are often placed in a single classroom with a single teacher, limiting their opportunity to observe various teaching styles and strategies. Although technology used to facilitate communication in a Type I field-based experience (e.g., online discussion forums) may allow for more sharing about students' various field placement environments, the true benefit of being exposed to a variety of classroom
environments is primarily associated with Types II and III experiences where students have access to a range of virtual and/or real classroom environments.

Teachers using video-based observations often note that a key benefit is the opportunity for students to observe a variety of teaching strategies and assessment techniques (Green, 2005; Hannah & Abate, 1992–1993; Lambdin et al., 1996; Putnam & Borko, 2000). Additionally, with the help of technological tools, preservice teachers can interact with students in culturally diverse areas. For instance, Lehman and Richardson (2003) describe a virtual field experience which utilizes two-way video-conferencing to connect preservice teachers in mid-central Indiana with diverse populations in other areas of the state. Seventy-six percent of the students participating in the virtual field experience reported that they felt more comfortable in their ability to understand and teach diverse learners as a result of their virtual field experience.

Virtual classrooms can also offer an environment for experiencing various student characteristics and learning about special teaching approaches not easily observable in regular classrooms. In a study by Foley and McAllister (2005), preservice teachers who engaged in the Sim-School© reported that the complex diversity of the simulated environment in terms of racial, cultural, and language needs of virtual students was a useful context for them to become more aware of those issues, to understand students’ unique characteristics, and to plan for differentiated strategies. Compton, Davis, Graham and Swaharu (2008) reported a similar study of virtual field experiences where preservice teachers observed an exemplary Virtual Schooling (VS) teacher to learn about the special set of knowledge and skills necessary to teach in the VS environment.

**Creation of Shared Experiences**

Since students usually participate in a traditional field experience individually, it can be difficult to hold meaningful discussions when students have not had the same experiences. The benefits of having shared experiences are realized primarily in Types II and III experiences where students have direct access to common experiences which may be in a live classroom via remote observation, pre-recorded classrooms, and/or simulated classrooms. The increased communication and documentation of experiences through technologies used in Type I experiences may also provide benefits related to shared experiences.

Faculty identify the creation of a common experience that can be analyzed and discussed by a group of students as one of the key benefits of using video-based cases over the traditional field experience (Green, 2005; Hannah, 1995; Hannah & Abate, 1992–1993; Rosen, 2003). Kent (2007) argues that from instructors’ perspectives, implementing interactive videoconferencing to observe real teaching strategies in a remote classroom is a practical way to ensure that preservice teachers are internalizing and emulating model teaching methods that educators set as standards. Students participating in a virtual classroom observation also recognize the value of having a shared experience which allows them to “check their perceptions of the class with other students and the instructor” (Knight et al., 2004, p. 146). A web-based peer assessment system developed by Wu and Kao (2008) promotes interaction among students by allowing them to view field-teaching sessions of their classmates and provide constructive feedback to each other.

The use of commonly viewed classrooms also offers an element of quality control in that the education faculty member is at least observing the same classroom (in case of two-way video) or perhaps selecting appropriate classrooms/cases (in case of pre-recorded video-based classrooms) (Hannah & Abate, 1992–1993; McDevitt, 1996). Using online discussions to share individual student experiences and hear different interpretations of the same event is also noted as a key benefit of technology-enhanced field experiences (e.g., Kale, et al, 2006; Moffett, 2001)

**Promoting Reflectivity**

Many researchers have examined the use of online technologies to encourage students to actively process and reflect upon their experiences in the field (e.g., Bonk et al., 1998; Edens, 2000; Mason, 2000; Meyers, 2006; Moffett, 2001; Whipp, 2003). The use of various forms of technology to promote critical reflection can be realized in different ways and to different degrees in each type of technology-enhanced field experience. Incorporating technologies that promote communication and interaction between peers and/or supervisors can enhance reflectivity in all types of
experiences. The use of video (Type II) and simulations (Type III) can offer other benefits that further promote critical reflection.

Video technology can be used to promote reflective practices and focused observations. Stockero (2008) found that preservice teachers who used a video case-based curriculum engaged in deeper reflection more specifically focused on student thinking. The video cases seem to promote an "ability to recognize subtle differences in student thinking by providing an opportunity to compare student thinking both across and within the cases" (p. 504). Similarly, Rosaen, Lundeberg, Cooper, Fritzen, and Terpstra (2008) found that video-supported reflection, compared to memory-based reflection, enabled preservice teachers to shift their focus of reflection from superficial features of classroom management to pedagogical issues.

The use of simulations in Type III experiences also offers benefits related to reflective practices by providing an error-free safe environment where preservice teachers can experiment with a variety of teaching strategies, such as on-the-spot-experiments, suggested by Schön (1987) as being important for developing the skills of a reflective practitioner. Payr (2005) argues that virtual field experiences reduce the complexity of the learning environment and present simplified activities that allow student teachers to engage in focused experiences without the flow of complex and irrelevant information interrupting the process of observations.

Preparing Students Cognitively

Many researchers have investigated the possibility of using video-based classrooms in addition to or instead of early field placements in order to better prepare students for future field placements. The use of video-based classrooms in Type II experiences and simulations in Type III experiences can help students to gain more from classroom-based experiences (Type I or traditional field experiences) they will participate in at all stages of their teacher education program.

Hannah (1995) emphasizes the importance of providing a “safe environment where students can explore their beliefs about teaching” and be introduced to observations in the field before “their upcoming foray into the real world of classrooms” (pp. 276–277). There is evidence to suggest that the use of technology related to field experiences creates such an environment where students are able to develop observation skills with the guidance of a faculty member who is observing along with them (e.g., Hannah & Abate, 1992–1993; Lehmnan & Richardson, 2003; McIntyre & Pape, 1993). Brooks and Kopp (1990) suggest that an interactive videodisk program counters negative aspects of field observation, particularly “untrained eyes and selective memories” that permit novices to “see what they want to see” (p. 500). This is consistent with the more recent research that has emphasizes the role of video for guiding explicit and focused observation skills (e.g., Brophy, 2004; Calandra, Gurvitch, & Lund, 2008; Sherin & van Es, 2005).

When preservice teachers are trained to learn from their early field experiences through the mediation of technology, they can fully benefit from their subsequent field-based experiences in classroom settings. Santagata et al. (2007) found that the use of virtual-video-based field experiences helps preservice teachers move from simple descriptions of classroom actions to focused observations about student learning and teacher interaction, and subsequently apply these skills in authentic settings. In another study, Baker (2005) found that preservice teachers who participated in multimedia case-based instruction (M-CBI) reported that the "analyses, reflections and discussion of M-CBI made their field experiences richer" (p. 428).

Learning about Technology Integration

It is obvious that technology will continue to have a large impact on the future of education, and should therefore play a prominent role in teacher education. However, it has been reported that preservice teachers often fail to integrate technology in field experiences due to the situational complexity and pedagogical difficulties (Dawson & Dana, 2007). Utilizing technology in relation to field experiences provides another opportunity for prospective teachers to become familiar with technology and understand its potential application in their future classrooms (Lehmman & Richardson, 2003). All types of technology-enhanced field experiences (Types I, II, and III) expose students to different types of technology and allow them to see how they can be used to enhance learning
experiences. For instance, Mason (2000) found that experiences with technological tools in technology-enhanced field experiences not only encouraged student teachers to integrate technology in their teaching practices, but also changed their perceptions and attitude toward technology use.

Limitations of Technology-Enhanced/Virtual Field Experiences

Because published articles tend to focus on the positive impacts and roles of technology, challenges and issues arising from the experiences of integrating technology-enhanced field experiences are not often reported or discussed in detail. However, it should be noted that while there are many advantages to using various forms of technology in relation to field experiences, there are also some serious concerns that must be addressed. This section presents four potential limitations: (a) lack of interaction with teachers and students, (b) limited reality and complexity, (c) availability of relevant cases, and (d) technical problems.

First, preservice teachers may view technology-enhanced or virtual field experiences as a lack of opportunities to observe real classrooms and interact with students in classroom settings. Knight and colleagues (2004) reported that although students did see value in their participation in a virtual field experience, they were troubled by the limited interaction capabilities of observing in this manner. Students noted that they were not able to interact with the students as they likely would have been able to do in a "live" observation, and some felt that the experience “depersonalized” the students (p. 147). Students were also concerned about the authenticity of the classroom they were observing, fearing that the “intrusive nature of the cameras” may have altered both student and teacher behaviors (p. 146).

Other researchers have reported that students who participate in a Type II or III experience may resent the virtual experience and view it as a missed opportunity to participate in a more authentic setting (Atkins, 1998; Lambdin, et al., 1996; McDevitt, 1996). In a virtual field experience described by Malewski and colleagues (2005), the impersonal nature and limited interaction of such experiences was addressed by adding two site visits to the school they were observing via two-way videoconferencing. A site visit at the beginning of the experience gave preservice teachers the opportunity to interact with and get to know the students and host teacher. A site visit at the conclusion of the activity “provided preservice teachers with a sense of closure” (p. 425).

Secondly, another issue is related to the limited reality and reduced complexity in vicarious or virtual classroom environments (Types II and III). In a review of the use of simulated classrooms in teacher education, Brown (1999) expressed that because the tradition of real classroom-based field experiences has been firmly established, it is difficult for students and teachers to see the value of technology-based field experiences as alternative activities, and they tend to perceive these experiences as being not real enough. For instance, Sharpe and associates (2003) examined the role of a videoconferencing tool to guide and supervise student teachers placed in different schools in Singapore. Participants were required to prepare and stream a short video clip demonstrating a particular teaching competency for discussion during the conference time. Participants reported the artificiality of video clips was a disadvantage since most of them tended to document their best teaching competency, which reduced the opportunity to observe and discuss several problematic issues in real classrooms. Reduced complexity may be beneficial for the purpose of teaching particular strategies, but the complex and ill-structured nature of classroom situations should be included to guide and teach other sets of skills and knowledge (Putnam & Borko, 2000).

Thirdly, an additional concern is related to the extent to which the video and virtual classrooms reflect students’ teaching interests. When placed for a traditional field experience, students are often placed with teachers of the grade level and/or subject focus they are interested in pursuing (e.g., a music education major wanting to teach elementary students would be placed with an elementary music teacher). When students are not able to view classrooms that they see as representative of their future classroom, they may view the video cases as irrelevant and fail to see how they apply to their educational context (Ma, Lai, Williams, Prejean, & Ford, 2008; Pape & McIntyre, 1993).

Lastly, technical problems may negatively affect perceptions about technology-integrated field experiences. O’Connor, Good and Greene (2006) reported a case study that explored the use of teleobservation through a videoconferencing tool to enhance preservice teachers’ observation skills prior to their final internship. While participants saw value in the teleobservation, over 95% of them reported technical problems with the quality of audio and video technology. Rhine and Bryant (2007) identify several concerns that need to be addressed when
implementing a technology-enhanced field experience, including the accessibility of necessary technology tools, the diverse range of technical skills among student teachers, and their comfort levels with the use of online discussion forums. While Rhine and Bryant attempted to address these issues, they found that student teachers’ reactions to technology-enhanced field experiences were fairly consistent with their attitude toward and skill levels related to technology. Overall, these studies imply that technology may become a barrier when issues surrounding the use of technological tools are not appropriately addressed.

These are legitimate concerns that must be considered when technology is used to supplement or replace a traditional field experience. While the technology may offer some advantages (as suggested above), it may also present new challenges. The benefits and limitations of using technology related to field experiences must be weighed carefully.

Discussions and Conclusion

It should be noted that the purpose of this paper is not to suggest that technology-enhanced/virtual field experiences are more effective than traditional field-based experiences. Indeed, many research studies reviewed in this paper suggest that technology alone does not make field experiences effective, but pedagogical considerations are critical to the design and implementation of successful technology-integrated field experiences that meet intended goals. Thus, when alternative field experience formats are being considered, strengths and weaknesses of each of format summarized in Table 1 must be weighed equally.

### Table 1. Benefits and limitations of traditional and technology-enhanced/virtual field experiences

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td></td>
</tr>
<tr>
<td>Face-to-face interaction with students and teachers</td>
<td>Time and travel constraints</td>
</tr>
<tr>
<td>Apprenticeship</td>
<td>Supervision of remote students</td>
</tr>
<tr>
<td>High reality and complexity</td>
<td>Access to rural and diverse settings</td>
</tr>
<tr>
<td>Technology-enhanced / Virtual</td>
<td></td>
</tr>
<tr>
<td>Exposure to various teaching/learning styles</td>
<td>Lack of interaction with teachers and students</td>
</tr>
<tr>
<td>Creation of shared experiences</td>
<td>Limited reality and complexity</td>
</tr>
<tr>
<td>Promoting reflectivity</td>
<td>Availability of relevant cases</td>
</tr>
<tr>
<td>Preparing students cognitively</td>
<td>Technical problems</td>
</tr>
<tr>
<td>Learning about technology integration</td>
<td></td>
</tr>
</tbody>
</table>

As is documented in the literature, field experiences can serve a number of purposes ranging from exploring teaching as a career option to linking theory with practice to developing skills necessary to be an educator (e.g., Aiken & Day, 1999; Hopkins, 1995). The overall goals of each specific field experience should be the cornerstone of all field experience format decisions. The technology-enhanced or virtual field experience options may be more effective in meeting some field experience objectives than others, and the same is true for more traditional field experiences. Let us consider some specific purposes of field experiences in relation to each type of field experience format.

One purpose of field experiences identified by Paese (1996) is to explore teaching as a career option and practice the teaching skills needed for carrying out that professional role. If this is the primary goal of a field experience, then it is necessary to ensure that the field experience format allows for interaction and dialogue with classroom teachers to learn more about teaching as a career. While it may be possible to construct a virtual field experience that allows for such opportunities (interactive Type II or III experience), special attention should be paid to common limitations identified by those participating in virtual field experiences – limited interaction and impersonal nature of the interactions (e.g., Knight et al., 2004; McDevitt, 1996). Similarly, the experience would need to provide opportunities for preservice teachers to actually teach students so they can develop the required skills. A study conducted by Atkin (1998), found that many students felt the lack of a traditional field experience actually impaired their ability to develop the required content knowledge. They felt that they needed the opportunity to teach students and test out their ideas with them in order to fully develop the necessarily knowledge and skills. It would seem that these goals would be more directly and effectively met through a traditional or Type I technology-enhanced field
experience where preservice teachers are able to experience a “real” classroom first-hand and interact with both teacher and students.

If the focus of a field experience is for students to decide if teaching is an appropriate career choice and/or decide upon a certification area (as suggested by Aiken and Day, 1999), it is critical to ensure exposure to a variety of classrooms. Technology may offer a way for students to easily observe a range of grade levels and subjects, thereby providing input for making such decisions. However, additional research would need to be conducted to determine if the somewhat limited observation that happens in a non-live setting (as in Type II experiences) would be sufficient for helping students truly get a sense of what teaching entails and what grade level/subject they would be best suited for.

Another common goal of field experiences is to help students better understand the conceptual and theoretical knowledge being presented in their college courses (Frey, 2008). There are many aspects of technology-enhanced/virtual field experiences that support students’ ability to see the theories they are learning about in practice. Some benefits technology offers in the Types II and III experiences are the creation of a shared experience that preservice teachers and their instructors can look at and analyze together; the ability to easily share experiences and reflections and receive valuable feedback; the exposure to a variety of teaching styles and techniques that can be purposefully selected or identified by an instructor; and, the opportunity to review classroom events multiple times (assuming the classroom is recorded in some format) to further analyze events and make connections to educational theories and concepts. Technology may be a useful aid in helping students better understand how theories are put into practice.

Traditional approaches for field experiences have been criticized for promoting utilitarian perspectives and uncritical acceptance of the current teaching practices (e.g., Beyer, 1984; Goodman, 1985, 1986). Technology may be a viable option to increase access to quality classrooms embodying types of pedagogical practices consistent with educational reform, and to encourage preservice teachers to explore new ideas in a safe environment. It is hoped that the information provided here will be of use to individuals who are considering technology-enhanced or virtual field experiences as an alternative to the traditional field experience. It is clear that the overall goals and objectives for a specific field experience must be the focus when field experience options are being explored.

References


The STIN in the Tale: A Socio-technical Interaction Perspective on Networked Learning

Steve Walker¹ and Linda Creanor²
¹Faculty of Mathematics, Computing and Society, The Open University // s.walker@open.ac.uk
²The Caledonian Academy, Glasgow Caledonian University // l.creanor@gcal.ac.uk

ABSTRACT

In this paper, we go beyond what have been described as ‘mechanistic’ accounts of e-learning to explore the complexity of relationships between people and technology as encountered in cases of networked learning. We introduce from the social informatics literature the concept of sociotechnical interaction networks which focus on the interplay between participants, technology, learning artefacts and practices. We apply this concept to case material drawn from transnational trade union education to identify and to analyse three aspects of networked learning: the local sociotechnical networks of learners; the construction of an overarching, global sociotechnical network for learning; and the evolution of such networks over time. Finally we identify issues for further research highlighted by these models.

Keywords

Networked learning, Social informatics, Social action, Web 2.0

Introduction

The learning landscape is becoming increasingly complex. In part this is due to technological developments such as Web 2.0, virtual worlds and social networking practices, alongside other factors such as widespread national and international policy changes, changing workplace skill requirements and growing learner demand for flexible learning arrangements. To understand this complexity, educators in diverse contexts require models and concepts which can help to make sense of, and to capitalise on, the interplay between people, technology, learning artefacts and learning processes. Networked learning provides a useful framework which encompasses not only pedagogy, but also the broader social, technical and cultural forces at play (Jones, 2004). The network metaphor which Jones describes as, “... a unifying concept allowing us to bring together apparently disparate elements of the field” (p81) remains compelling. It underpins our thinking as we explore models of networked learning which extend beyond the traditional confines of the formal tertiary education sector.

Designing and participating in effective networked learning are significant accomplishments in which educators and educational technologists ‘orchestrate’ groups of people using technology, tailored learning activities and a range of learning resources to enable learning (e.g. Barab et al, 2004; Walker & Creanor, 2005). Educators frequently need to integrate practices associated with networked learning alongside those of face-to-face learning, creating a ‘blended’ environment. Learners participate in individual and collaborative activities through which they can develop new meanings, skills and knowledge. In doing so they may use technologies which are new to them or, increasingly, they may be integrating their e-learning activities into an ‘underworld’ of communication through personal social and mobile technologies, of which tutors may be unaware. (Creanor et al, 2008). In this paper we draw on the social informatics and sociotechnical traditions of research into information and communications technologies (ICT) to highlight the complexity of interactions between people and technology in networked learning situations, and the consequent potential sensitivity to apparently trivial difficulties. Our primary contribution is to demonstrate how a particular approach from these traditions, the sociotechnical interaction network or STIN (Kling et al, 2003; Scacchi, 2005; Meyer, 2006) can be used to think about these complex interactions.

Sociotechnical studies have established that technology design and use are complex outcomes of multiple, interacting influences operating at different levels. At the micro-level for example, organisational and social influences include incentive structures and local working or learning cultures (e.g. Orlikowski, 1993; Kling, 2000). Simultaneously at macro-level, influencing factors identified by, among others, Agre (1998) and Williams (2000), include social (e.g. in the way particular communities of practice are embedded within institutions), political (e.g. how technologies may be promoted or regulated) and economic (e.g. the various cost factors associated with particular types of transaction). Studies of networked learning in these traditions (e.g. Hara & Kling, 1999; Kling & Courtright, 2004; Dutton et al, 2004) offer an alternative approach to understanding the evolving relationship of learning and technology to those
which comprise what Diana Laurillard has referred to as a historically dominant ‘mechanistic’ account of change (Nash et al, 2004; Laurillard, 2005). The weaknesses of mechanistic models can be seen in discrepancies between claims made about learning technologies and the reality of their use. These have been demonstrated both at the micro-level of student responses to technology enhanced learning (e.g. Hara & Kling, 1999; Sharpe et al, 2005; Creanor et al, 2008) and at wider institutional and political levels (Selwyn, 2007).

Sociotechnical studies have generated a rich collection of methods, concepts and findings about how technology is implemented and used. In the following sections we use and develop one of these, the concept of the STIN, originally proposed by Kling et al (2003), to analyse the complexity of three aspects of networked learning. Drawing on case material from a transnational trade union education initiative, we distinguish the concepts of ‘ego-STIN’ (a network viewed from a particular individual perspective) and ‘whole-network STIN’ (an overarching network which encompasses a number of ‘ego-STINs’). In the subsequent discussion we conclude that while the STIN concept offers a potentially fruitful approach to considering the complexity of real-life networked learning, further work is needed, in particular to follow changes in the configurations of STINs over time.

Framework: sociotechnical interaction networks

Network models have been widely used to capture the complexity of relationships between people and technology at multiple levels of analysis from the macro levels of social forces (e.g. Law and Callon, 1992) to the micro level of particular technologies (e.g. Kling & Courtright, 2004). In these models, technology supported networks are viewed as collections of artefacts and people linked in multiple ways by practices, protocols and understandings. These models reflect the complexity of introducing new artefacts into existing networks, where outcomes are frequently unpredictable and may propagate through wider networks to have effects often far removed from the original intentions.

The sociotechnical interaction network (STIN) is one such networked model, in which the technological is seen as co-constitutive with the social. Kling (2000) refers to such a network as ‘highly intertwined’, in that the technological elements cannot sensibly be discussed independently of the social aspects. This view does not, however, ‘...insist that this intertwining of technical and social elements is universal. Rather, it is commonplace, and a good heuristic for inquiry, especially with complex technologies’ (Kling, 2000a:220).

Behaviour is thus not simply dictated by the affordances of a particular technology or artefact, but through participants interacting with both people and artefacts which may themselves also be part of other networks. Kling et al (2003) illustrate this approach through an example of online academic communication in which they distinguish several conceptual differences from what they call the ‘standard model’ of technology use. Firstly, their analytic focus is ecological, deliberately looking beyond the affordances of the technology or the narrow relationships between participants and artefacts in a particular network. Secondly, a limited view of the ‘user’ is replaced with a wider view of participants as social actors (see also Lamb & Kling, 2003) who have multiple roles and relationships which can affect behaviour in a STIN under analysis by linking that STIN to other relevant STINs in multiple ways. Thirdly, technology is viewed as open to local adaptation and social influence (‘configurational’), rather than simply offering a limited set of functions. Kling et al (2003) argue that this approach is better able to capture the complex interaction between the digital systems and the ‘real world’ institutions and practices of academic life.

In using the STIN approach the scope and appropriate level of detail to be included in the network are determined by the researcher, relative to the issue under analysis: different levels of resolution will be appropriate to different analyses. STINs are recursively embedded within each other such that it is always, in principle, possible to break down individual elements of STINs further, to reveal the networks within them. STINs, then, offer rich models of complex social, political, economic and in the cases discussed here, pedagogic, interactions and can be used at multiple levels of analysis.

In developing our analysis of the relationship between technology and learning, we have chosen to concentrate on a micro-level analysis of specific arrangements of people, technologies and practices. We use diagrams of STINs to visualise and illustrate the important social and technical nodes and diverse links between them. Comparing the configuration of these networks then allows us to illustrate commonalities and differences between the heterogeneous configurations of actors within them. It is in this metaphoric mode that we use the concept of STINs.
to describe and analyse three aspects of networked learning: the relationships within ‘local’ configurations; between ‘local’ and ‘global’ configurations’; and as a way of modelling the way these configurations may change over time. Below we reconceptualise these three issues in related literatures as sociotechnical interaction network terms before testing these models against case study materials.

Local configurations

Participants in networked learning commonly come from a diverse range of backgrounds where they are embedded in a range of occupational, domestic or other sociotechnical networks and practices. These have important implications for how they gain access to the resources they need to participate in networked learning. Representing participants’ local situations as STINs allows us to identify important elements in a particular setting, and the relationships and interactions between them.

Our case material is drawn from examples of e-learning in a trade union setting, where learning events are conducted largely outside formal educational institutions, and in which informal and peer elements of learning and access to resources play a significant role. Sawchuk (2003) has demonstrated how resources are mobilised through social networks to enable learning about technology in working class communities in Canada. These informal ‘working class computer learning networks’ are central to the development of their members’ knowledge about computers. Sawchuk’s identification of these networks suggests that participation in networked learning may similarly rely on available social capital, understood as the ability to mobilise a variety of resources (including information) through social networks.

We can use STINs to illustrate these local circumstances, using them in ways analogous to the ‘ego’ network of social network analysis (Wasserman & Faust, 1994:41) which traces the network links to and from an individual person. Ego network studies have been widely used in disciplines such as social anthropology and clinical psychology to examine the sources of support available to individuals or families. While in a STIN, these ‘ego’ networks will be heterogeneous, made up of the social and technical practices associated with each learner, they allow us to identify the support available to participants, and barriers to participation in wider sociotechnical networks. We develop this argument in the first case study below.

Networked learning as the relationship of local and global configurations

If, as in the previous discussion, we conceive of participants’ local contexts as STINs, then we can think of the design of networked learning as linking together these local STINs to enable learning; a networked learning event becomes a ‘network of networks’. The design of such an event becomes an exercise in ‘heterogeneous engineering’, bringing together people and technologies organised through pedagogic practices and artefacts.

When we think of a networked learning event as a STIN, we are interested in the totality of interactions between the people and artefacts involved, and the subsequent outcomes for learning. Rather than placing an individual at the centre of the network, as in the local ‘ego’ STIN, we are interested in the structure and properties of the whole network; Kling et al (2003) work at a whole-network level of analysis in their study of scholarly communication. The social network analysis literature (e.g. Wasserman & Faust, 1994) suggests a rich range of concepts which can help us to think about properties of such STINs. These include, for example, network centrality, structural equivalence, and weak ties. In the networked learning literature, discontinuities between participants’ local STINs can be thought of as boundaries, which, given the heterogeneous nature of STINs, can take multiple forms (e.g. participants may be using different software; they may come from different social settings; they may have different mother tongues). We have previously argued that networked learning interventions can be thought of as boundary encounters, and that depending on the nature of the intervention, boundaries may either be central opportunities for, or obstacles to, learning (Walker & Creanor, 2005). In such settings, facilitators play a particularly important role in identifying potential cross-boundary interactions and designing appropriate pedagogical artefacts and collaborative activities. We explore this further in the second case study below.
Changing configurations over time

The structures of many STINs are likely to be time sensitive. Barabasi (2002) has demonstrated that network structures vary depending on the rules governing how new nodes are joined to a network. While the nature of nodes and the links which connect them are much simpler in Barabasi’s examples (e.g. networks of hyperlinks between web sites) than in the examples with which we are concerned, we may similarly find that the manner in which STINs emerge or are designed influence their structures and properties in important ways.

In our case examples, the sociotechnical configurations of networked learning may change over time at three levels. Firstly, an individual’s ‘ego’ STIN may be reconfigured by their participation in a learning event and the consequent development of new social relations or the construction of new knowledge or skills. For example, new technology-related skills may result in the reconfiguration of some of the technical elements in the ‘ego’ STIN. More widely, as people’s informational needs evolve through different stages of their careers, they mobilise information and learning in different ways (Penuel & Cohen, 2003), a process which may be amenable to analysis as a reconfiguration of ‘ego’ STINs. Secondly, within a learning event different sociotechnical configurations may be appropriate to different stages of, or activities within, a learning process. The learning design may incorporate changing activities, social relations and uses of technologies during learning intervention. Thirdly, particularly in social action settings, an explicit aim of networked learning may be to achieve some form of longer term social change. Learning interventions may be aimed either at equipping individuals to participate more effectively in wider social activities or at supporting interactions among participants beyond the life of the learning activity itself. This involves moving from learning together, to doing or working together.

Exploring networked learning therefore may contribute to the emerging understandings of how virtual communities change over time (Andriessen, 2005) and how learning, in these less formal guises, can be supported.

Case studies

Our case studies come from trade union education, where learning is frequently conducted outside formal education contexts. It is also characterised by the values informing the design and content of the learning and the intent to encourage some form of collective social change. Pedagogies are varied, but in general tend to be learner-, rather than teacher- or expert-centred, often using co-operative or collaborative methods. The collaborative dimension of the learning is a part of the process of social change, and in some cases is intended directly to encourage collaborative working beyond a particular learning intervention. The audience for trade union learning is often comprised of people without histories of engagement with educational institutions, and in some cases with active alienation from them. It can, as in the examples discussed below, involve bringing together people who are living and working in very diverse settings.

In the following section the three uses of STINs outlined above will be illustrated by drawing on case material from Dialog On, a collaborative 16-partner project supported by the European Social Fund (ESF) and led by the European Trade Union College (ETUCO). The project took a networked learning approach to building the capacity of unions to organise in rapidly changing economic circumstances. The project was organised in two strands: a computer-mediated distance learning (CMDL) strand and a ‘networking’ strand. In the CMDL strand, experienced trade union educators were trained by a team of educators knowledgeable about e-learning methods, with academic support, to design and deliver fourteen national and transnational blended mode courses. The transnational courses were organised as two residential workshops with intervening periods of around three months of online learning activity. The topics of the courses were varied, for example, focussing on ‘Regulation of atypical employment’, or ‘Competencies for negotiating on issues of vocational training’. The target audiences were correspondingly varied, though the transnational courses all involved trade unionists from pairs of European Union countries.

In the networking strand trade unionists from particular industrial sectors (e.g. the graphical industries, or higher education) participated in eight learning networks organised by European sectoral trade union federations to exchange information and generate knowledge about developments in their sector. The networking activities were similarly organised as combinations of residential workshops and facilitated online activities. Pairs of ‘animateurs’ were trained for each network, one of whom specialised in operational and pedagogic aspects of the network, and the other in organisational aspects. Most of the networks were intended to become self-sustaining, beyond the life of the
Both strands were supported by centrally-produced materials and a common communication infrastructure (the First Class conferencing system). A total of 27 CMDL strand tutors and animateurs were trained in online learning and network facilitation methods before embarking on their own courses or networks. Over the two years of the project, tutors delivered 32 distance learning courses to 471 trade union learners, while the 8 online industrial sector networks attracted over 300 participants. The data used below were derived primarily from the project evaluation activities.

Each transnational course was treated as an exploratory case study of a real world event (Yin, 2003) seeking to identify constraints and benefits of transnational e-learning in trade union education. Each case combined quantitative and qualitative data. Course participants completed an initial self-profile questionnaire covering age, gender, knowledge of languages (spoken and written), union position and responsibilities, prior experience of transnational union collaboration, and prior experience of both trade union and distance education. At the end of each course, participants were asked to complete an evaluative questionnaire, covering their views on the course, materials, expectations of technology, their experiences of online collaboration and their expectations of their ability to apply their learning, all on 5-point Likert-type scales. The questionnaire also included open-ended questions asking for more detailed views of what they had liked about the course and any suggestions for improvement. The online activities were analysed quantitatively, with information collected on distribution of conference contributions temporally, by participant and within sub-conferences (usually corresponding either to particular sub-tasks or work allocated to sub-groups in the design of the courses). Course tutors were asked to complete an evaluation questionnaire, presented their experiences at an end-of-project conference and discussed issues raised at an evaluation workshop designed and attended by the authors.

The data were assembled into narratives summarising the conduct of each course – in particular patterns of online conference usage, the relationship to the learning programmes (as designed by the course tutors and network animateurs), and both tutor and participant comments on the courses, the achievements, and the difficulties encountered. In the examples below we have used the construct of the sociotechnical interaction network to consider instances of difficulties encountered by tutors and/or participants. These failures, we suggest, highlight aspects of sociotechnical interaction networks which otherwise remain unremarked, or even unnoticed.

**Local sociotechnical interaction networks**

Dialog On was concerned with enabling participation in networked learning by trade union members living and working in very different situations. Some elements of these situations were common across groups or sub-groups of participants but some were unique to an individual.

Participation in networked learning events requires access to a communication infrastructure. Access, however, is embedded in a range of organisational and domestic circumstances. Some of these relationships come to the fore clearly in the case of trade union education, where participants may, for example, be full-time union employees participating as part of their paid employment, workplace representatives with an office and technology access provided by agreement with (or legal requirement on) the employer, or activists with no access in the workplace but with access from home. Domestic and office settings have differing enabling and constraining implications for access. Below, we present these differing situations as micro-level sociotechnical interaction networks. We use the example of firewall restrictions on access to the project conferencing server. This was the most widely reported technical problem in the project, occurring frequently in participant course evaluations and interviews with network animateurs and course tutors.

The project conference system could be accessed either via a web interface or by a dedicated software client downloaded to the user’s PC and which communicated with the server using a proprietary protocol. Most participants were trained in the use of the conference system in project workshops. The training emphasised the client rather than the web interface, partly because some functionality was only available via the client and partly because, once learned, the client interface was thought to be easier to use. However, a widely reported problem across the project was that of using the client to access the server through organisational firewalls, a problem which derived from the client’s use of seldom-used Internet connection settings. For example:

"Some problem with "firewalls" obliges me to go through the Internet... and it is slower..." –Participant evaluation questionnaire
A more complex picture emerged during discussions by network animateurs at a mid-project evaluation event. For security reasons, many network managers set firewalls to block all connection settings except those explicitly permitted, for example to support applications such as web access or email. A frequently reported experience, following the training, was of participants returning to their organisations and finding that, having installed the client software as instructed, they were unable to connect to the server. Participants were advised to discuss their problem with whomever was responsible for network security in their organisation. While many network managers were responsive to the problem and ‘opened’ the firewall to client traffic, others were not. In such cases, despite the training, participants were forced to use the web interface. However, the problem recurred on several occasions even where the firewall had been opened: in instances where network managers opened the port informally, the new settings were lost when a firewall was upgraded. From a user perspective the firewall appeared arbitrarily to deny access once more. For these participants, achieving and maintaining access to the server was a personal and organisational accomplishment as much as a technological one (Figure 1).

![Figure 1: Accessing external conference server from within an organisation: stylised example of an ‘ego STIN’](image1)

For those participants who accessed the project server from the home, a rather simpler ‘ego STIN’ (Figure 2) illustrates the organisationally simpler environment.

![Figure 2: Accessing external conference server from home: stylised example](image2)
Given the difficulties in accessing the course infrastructure, there was an immediate danger of a negative impact on participants’ motivation and their subsequent level of participation in the course. Learner motivation can be affected by many factors (e.g. Warren, 2000), but for networked learning in particular, ease of access is fundamental. By capturing aspects of social and organisational arrangements, the STINs help us to consider what is frequently conceived of as a simple issue of technical access as something rather more complex.

**Networked learning events as ‘global’ networks**

The CMDL courses were designed to bring together trade unionists from pairs of countries to study some aspect of industrial change, thus encouraging the development of a wider European perspective. Conceived of as sociotechnical interaction networks, they formed a network of local STINs for the duration of the course. One of the aims of the courses was to enable learning from collaborating with trade unionists from other countries. This is a particular challenge for European trade unionists, where industrial relations systems, trade union organisation and ways of working vary radically from country to country. As some aspects of workplace regulation are now agreed at the level of the European Union and with progressively more transnationally integrated work methods, trade unionists need increasingly to work with others in very different situations.

The course considered here brought together 16 experienced French and Spanish trade unionists in a blended mode course of two residential workshops and an intervening period of 18 weeks of online small-group collaborative learning activities. These online activities involved four working groups (two of Spanish participants, two of French participants) preparing presentations on topics identified at the first workshop for discussion at the second. As well as collaborating in their own small groups, participants were also encouraged by tutors to share their progress with the other working groups. The online part of the course relied on the First Class conferencing system, configured to support the planned working patterns for the learning tasks. The course can be thought of as an attempt to create a global sociotechnical interaction network which brought together participants operating in their own diverse local STINs.

![Figure 3: Sociotechnical interaction network diagram of distance learning case study](image_url)

The communication practices during the course reveal the importance of these local sociotechnical networks. In this case, the problems with building bridges between the local networks during the distance phase demonstrate the difficulties encountered in establishing an effective ‘network of networks’. Participants from each country shared linguistic, organisational and, to some extent, geographic, commonalities. Linguistic commonalities were reflected in
the design of the online tasks: the four working groups were each monolingual – two working in Spanish and two in French. In practice, however, the working patterns of the two language groups diverged significantly, with almost no interaction between them. According to an interview with project staff the French participants did not use the conference server as their primary communications medium, but reverted to more familiar email systems for their group work, effectively rendering it invisible to tutors and the Spanish participants. Part of the reason for this was that the technical training planned for the first residential workshop did not happen due to unforeseen contingencies. In contrast, the Spanish participants who were already familiar with the conferencing system from its use in their own confederations, used it in a way closer to the tutors’ original expectations. Additionally, many of the Spanish participants were based in Madrid and organised their own informal face to face meetings. Consequently, levels of electronic communications, as seen in usage of the conference server, were modest, despite encouraging support by one of the tutors. As described above, the Spanish group augmented their online working with local face to face meetings, while the French group used a completely different (and to the Spanish, invisible) communications medium. Figure 3 represents the observed interactions qualitatively as a STIN to highlight the discontinuities in communications between national groups, and the alternative communication channels which emerged.

Changing configurations of sociotechnical interactions over time

We expect many sociotechnical configurations to change over time. In Dialog On, participation in networked learning was aimed either directly or indirectly at improving the collective capacity of trade unions to respond to, and contribute to, shaping social change. In the CMDL strand this was largely indirect, through individual trade unionists developing knowledge and skills that would be of use to their trade union work. In the networking strand, the initial training interventions aimed to establish durable computer-mediated networks that were stable over time. The following case study highlights some of the challenges encountered.

The transnational network aimed initially to improve information collection on the state of collective bargaining in European countries in a traditionally well-unionised industrial sector. It was established by the relevant European sectoral federation and its affiliated national trade unions. Previously, this information had been gathered via an annual paper survey which, while effective, had required a great deal of administration. It also only gathered information ‘after the event’ of the various national and company level negotiations. By creating an online forum, it was expected that information could be shared more readily and that consequently network members would become ‘closer together’ and ‘more linked’, emerging as a durable learning and organisational network monitoring trends and developing the capacity to intervene. If successful, the network might also provide a model for other parts of the federation.

![Diagram: Collective bargaining information network initial configuration](image)

**Figure 4: Collective bargaining information network initial configuration**
The network was prepared at a residential workshop. Representatives from affiliated unions were to provide bargaining information using online questionnaires through a single central forum. Ten of the thirteen initial network members were employees of national unions, the remainder being workplace representatives. The network conferences were implemented against a background of informal, ad hoc email links among some of the participants, some of whom also met from time to time at meetings of the federation. The network was co-ordinated by a facilitator or ‘animateur’ as part of their responsibility working for the European Federation. The collation of collective bargaining information was initially handled by an academic with close links to the federation.

In the week immediately after the workshop, 27 messages were posted online indicating enthusiasm and engagement with the topic. Subsequently, this fell rapidly before growing again modestly. The animateur reorganised the network conference in week 18, creating a collection of six conferences with additional sub-conferences. In the five weeks before the reorganisation, a modest but consistent average of 7.5 messages were posted each week, demonstrating an ongoing interaction among network members. Afterwards, the average use (summed across all conferences) fell to less than two messages per week. Encouraged by early signs of growth, the animateur had tried to extend the range of the network, in part to encourage new participants to join. However, contributions became fragmented across locations, usage rapidly fell away, and the network ‘died’. The reorganisation, at best, appears to have been premature. The ‘before’ and ‘after’ states of the network are illustrated in Figures 4 and 5.

Discussion

We have argued that sociotechnical interaction networks are a fruitful way to conceive of networked learning interventions, and have used STINs illustratively to explore three elements of networked learning in social action settings. These case studies are suggestive of issues for further research in the rapidly changing networked learning environment, most particularly in relation to web 2.0 technologies. We then review some of the difficulties of using the STIN as an analytic device.

Networked learning and web 2.0

This fieldwork was conducted before the recent widespread uptake of web applications such as blogs, wikis, tagging, social networking, photo and video sharing, widely referred to as ‘web 2.0’. This growth has triggered extensive interest among educators both because of the changing expectations of technology use among learners, and the potential for new ways of organising learning episodes.
There is growing evidence (e.g. Prensky, 2006) that learners are approaching networked learning with complex computer-mediated communicational and informational networks already in place: they may be sharing photos and videos publicly and with friends and family, maintaining blogs, and developing a presence on social networking sites. These applications are prime examples of Kling’s ‘highly intertwined’ technologies, co-constituted by the technical and the social. Sites such as Wikipedia or Facebook by themselves do very little; it is the network of people, pages and practices which creates something of interest. For example, elaborate protocols and social structures have become established around Wikipedia to mediate contributors’ varying levels of interest, involvement and intent (e.g. Kittur & Kraut, 2008). Also, different social groups may be attracted to competing sites of the same general type: use of particular social networking sites appears to be influenced by factors such as ethnicity and family education (Hargittai, 2008). The technologies do not, alone, dictate who uses a site, how they relate to it, or, we may surmise, the meanings people attach to that use. Individual patterns of use of the same, or similar technologies, may vary radically. This may be more true than with previous online technologies, as sites and practices evolve differently over time. The variability of people’s use and understanding of technology may make it harder for designers of networked learning to make assumptions about learners’ attitudes towards and skills in the use of these technologies and their potential value for learning.

The STIN approach outlined in this paper offers a way of thinking about learners’ patterns of technology use in a way that can take explicit account of values, practices and motivations beyond concerns with skills or access. Considering users’ diverse local contexts as ego STINs allows the designer of networked learning interventions simultaneously to recognise the social and technical constituents of a learner’s participation. In turn, this has implications for the design of the wider (sociotechnical interaction) network of (sociotechnical interaction) networks that constitute a networked learning intervention. For want of a more elegant term, we will refer to this as an ‘interSTIN’. For example, rather than trying to encourage learners to interact with a central technical infrastructure which may fit more or less well into learners’ existing ego-STINs, we might take as an alternative starting point the question ‘what do we need to do to link up diverse ego-STINs in ways which encourage learning?’.

This might mean, for example, at a technical level creating access points to our ‘interSTINs’ through multiple social networking sites. It might also mean relinquishing tutor control and granting the learner a higher level of autonomy over a more personalised learning environment. It should also help us to guard against over-generalising what we think we know about how different groups of people use social technologies. We can be guided in these considerations by previous research in fields such as social informatics, computer mediated communications (CMC) and computer-supported collaborative working (CSCW) which have increasingly emphasised the important of understanding the relationships of the social and the technical. For example, longstanding findings on the significance of factors such as duplication of effort and status (Grudin, 1994) or critical mass (Markus, 1987), or more recent findings on performance in social networking sites (Liu, 2008) offer insights into the way the social and the technological interact in learners’ increasingly complex digital ‘ecologies’.

**Sociotechnical interaction networks**

While interest in the STIN approach appears to be growing, it remains at a rather undeveloped state of theoretical development. At this stage it is perhaps best thought of as an analytic strategy rather than a body of theory (Meyer, 2006). We argue that used in such a way, sociotechnical interactions have been useful tools for thinking about several aspects of networked learning. Our identification of the significance of ego-STINs and interSTINs may be a contribution to the development of a wider theory of sociotechnical interaction networks.

However, there are some weaknesses in the current state of the art. Firstly, the diagrams here are rather metaphoric, derived largely from the observations of difficulties encountered in networked learning events rather than successes; presenting other case studies in this way may allow us to recognise patterns associated with successful and unsuccessful interventions. One of our anonymous reviewers specifically questioned the need for diagrams at all, since the concepts were clear from the text. For us, however, the visual representations have been important tools in developing our analysis; while we have included them as part of our analytic strategy there is evident uncertainty about their value as a communication device. There may, though, be value in considering such sociotechnical networks more formally with network analysis tools, identifying more precisely the nature and relationships of nodes and the interactions between them. Kling et al (2003) assert that some STINs cannot be represented as directed graphs (a precondition for applying more formal methods); while there may be formidable practical difficulties in doing so, it is unclear to us why it is not possible in principle. Secondly, and more particularly, we have found that

314
network diagrams of the type we have used here do not adequately capture key elements of the temporal dimension of networked learning interventions without becoming overly complex. By building on more formal representations of STINs, it may be possible to capture the temporal dimension more usefully.

Conclusions

Our first foray into the use of sociotechnical interaction networks to help us to think about networked learning in social action settings has, we argue, been a useful way of conceiving of difficulties, and may help in designing alternative approaches to the organisation of such learning. We suspect that it will prove to be a powerful approach to analysing networked learning in the increasingly complex sociotechnical interaction networks of web 2.0 and social network technologies. Learners will be integrating planned e-learning activities into their complex individual information ecologies (JISC, 2007). There have been plenty of misleading predictions about how new technologies will render particular learning practices redundant. However the expectations of new generations of digitally literate learners may mean that there are new complexities to be incorporated into how we design our learning interventions.

We have also identified significant weaknesses in the practical use of STINs. In particular, the complexity of representations of changing configurations can make it difficult to identify and highlight the most significant changes. As with social network analysis, for STINs to be useful in dynamic settings, it may be necessary to simplify their representation through measurement of key features. It is, for us, an open question whether, given the heterogeneous collection of nodes and links they represent, STINs may usefully be open to more formal analysis using methods developed in, or analogous to, the analysis of other types of network (e.g. Barabás, 2002; Wasserman & Faust, 1994).

References


Developing Cognition with Collaborative Robotic Activities

Rubén Mitnik*, Miguel Nussbaum and Matías Recabarren

Department of Computer Science, School of Engineering, Pontificia Universidad Católica de Chile, P.O. Box 306 Santiago 22, Chile // Fax: +56-2-6864444 // rmitnik@ing.puc.cl // mn@ing.puc.cl // mrecabarren@ing.puc.cl

*Corresponding author

ABSTRACT
Cognition, faculty related to perception, imagination, memory, and problem solving, refers to internal mental processes through which sensorial input is acquired, elaborated, used, and stored. One of its importances relies on the fact that it affects in a direct way the learning potential. It has been shown that, even though cognitive processes develop side by side with biological maturity, this cognitive development can be enhanced by means of mediated learning as signaled by Feuerstein’s Mediated Learning theory. Based on this theory is that we propose an intervention model that addresses school academic issues using technologically assisted small group collaboration, pursuing a dual academic objective: to thrive students’ cognitive processes while addressing school curriculum topics. The purpose, therefore, is to balance the students’ cognitive differences by means of in-school content-filled classroom activities. Our aim is to make use of peer mediation in a real world setting with a virtual construction of it. In this paper, we describe this novel intervention model along with an in-school usage experience. For this, we present an activity designed for high school students, specifically aimed to assist the learning of kinematics, graph interpretation, and graph plotting. In this activity the students work in groups of three, using a robot and wirelessly interconnected Personal Digital Assistants (PDA). By means of a controlled experiment, we show how technologically-supported peer mediation promotes the students’ enrichment of their cognitive processes in each of the different stages of the mental act (input–elaboration–output), favoring communication skills, insight, and reasoning, while also restraining impulsive conduct and trial-and-error answers.

Keywords
Cognition development, Robotic assisted teaching, Interactive learning environments, Face-to-face computer supported collaborative learning, Intelligent tutoring systems

Introduction
Cognitive processes are the organizing mental structures which control the subject’s interpretation and storage of information (Neisser, 1967). This establishes, among others, the conduct toward stimuli, affecting in a direct way the capability of learning.

According to Piaget (1966), cognitive processes develop as a result of the interaction between stimuli and the organism. In this perspective, where cognitive development goes side by side with biological maturity, the organism is treated as a receptor of stimuli that, by means of these, builds up and matures its cognitive processes. Feuerstein expands on this idea indicating that human beings are cognitively modifiable (Feuerstein, 1970; Feuerstein et al, 1980). He proposes that the cognitive development of an individual can be enhanced by means of a mediator who selects, organizes, and structures stimuli in order to attain a specific goal. The results of the mediation are the acquisition, by the mediated subject, of appropriate behaviors, sets of knowledge, and operational structures which modify his cognitive structure, thus allowing the enrichment of his cognitive processes.

The aim of peer mediation is to develop mediation teaching styles and cognitive modifiability. Peer mediation is mainly based on Vygotsky's sociocultural and Feuerstein's mediated learning experience theories (Vigotsky, 1930/1978; Feuerstein, 1986). Experimental research has shown that peer mediation improves analogies capabilities (Tzuril & Shamir, 2007), thrives inventive thinking skills (Sokola et al. 2008), and increases scores in math (Shamir et al. 2006) and Science (Cattle & Howie 2008). There is also work performed in adults, where it has been used with organizational groups that need to be flexible enough to lead initiatives for change (Dawes, 2006).

Feuerstein’s Instrumental Enrichment Program (FIE) (Feuerstein et al. 1979) intends to modify the cognitive structure of disadvantaged students, or persons with special educational needs. This program, based on the Mediated Learning Experience (MLE) theory (Feuerstein, 1986), consists of a battery of activities, tests, situations, and
problems, constructed to modify the students’ cognitive deficiencies. Nevertheless, since the aim of the program is to correct cognitive functions, no specific academic issues are addressed.

In this paper we propose, based on Feuerstein’s theories but not on his instruments, an intervention model that addresses school academic issues using technologically assisted small group collaboration. We focus on a dual academic objective; to thrive students’ cognitive processes while addressing school curriculum topics. The purpose, therefore, is to balance the students’ cognitive differences by means of in-school content-filled classroom activities. Our aim is to make use of peer mediation in a real world setting with a virtual construction of it. We show our approach with high school children working in groups of three, using a robot, and each with a wirelessly interconnected Personal Digital Assistant (PDA). By taking advantage of the robot’s autonomy and accurate navigation systems, and the mobility of PDAs, we designed an activity designed to assist the learning of kinematics concepts, graph interpretation, and graph plotting. By means of a controlled experiment, we show how cognitive functions can be enhanced with technologically supported peer mediation.

This paper is structured as follows: Section 2 introduces the cognitive functions addressed; Section 3 illustrates the pedagogical model; Section 4 details the technological system; and section 5 describes the implemented activity and results. Finally, Section 6 presents the conclusions of this work.

Table 1. Feursteins defined cognitive functions

<table>
<thead>
<tr>
<th>Cognitive Function</th>
<th>Mental Stage</th>
<th>Ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear perception</td>
<td>Input</td>
<td>Accurately distinguish stimuli information</td>
</tr>
<tr>
<td>Precise information gathering</td>
<td>Input</td>
<td>Perceive information thoroughly, selecting relevant data in a care ful and organized way</td>
</tr>
<tr>
<td>Systematic exploration of a learning situation</td>
<td>Input</td>
<td>Organize and store the received information</td>
</tr>
<tr>
<td>Spatial organization</td>
<td>Input</td>
<td>Establish relations between events and objects located in space</td>
</tr>
<tr>
<td>Temporal orientation</td>
<td>Input</td>
<td>Establish relations between past and future events</td>
</tr>
<tr>
<td>Object conservation</td>
<td>Input</td>
<td>Preserve an object as invariant even though some of its attributes change</td>
</tr>
<tr>
<td>Organization of information</td>
<td>Input</td>
<td>Use different information sources, basic for establishing relations events and objects</td>
</tr>
<tr>
<td>Perception and definition of a problem</td>
<td>Elaboration</td>
<td>Determine what is asked by the problem, what data must be found, and how to find it</td>
</tr>
<tr>
<td>Selection of relevant knowledge</td>
<td>Elaboration</td>
<td>Choose relevant knowledge needed to solve the problem</td>
</tr>
<tr>
<td>Internal mental representation</td>
<td>Elaboration</td>
<td>Use abstract symbols and concepts to represent and generalize reality</td>
</tr>
<tr>
<td>Perceptual organization and structure</td>
<td>Elaboration</td>
<td>Direct, establish and project relations</td>
</tr>
<tr>
<td>Conduct planning</td>
<td>Elaboration</td>
<td>Develop a strategy which includes every step towards the goal or solution. Steps must have an order and a temporal sequence</td>
</tr>
<tr>
<td>Hypothetical thinking</td>
<td>Elaboration</td>
<td>Construct and validate hypotheses</td>
</tr>
<tr>
<td>Logic evidence</td>
<td>Elaboration</td>
<td>Prove answers by means of logic reasoning. Ability to justify a solution.</td>
</tr>
<tr>
<td>Explicit communication</td>
<td>Output</td>
<td>Use a clear and precise language to communicate an answer to another individual.</td>
</tr>
<tr>
<td>Uninhibited answer communication</td>
<td>Output</td>
<td>Express the answer in a fast, correct and systematic way</td>
</tr>
<tr>
<td>Answer Control</td>
<td>Output</td>
<td>Reason and consider before giving off an answer. Restrain trial and error kind of answers.</td>
</tr>
<tr>
<td>Accurate and precise answer</td>
<td>Output</td>
<td>Accurately identify and use different dimensions, such as size, distance and time, as well as relations among them</td>
</tr>
</tbody>
</table>
Cognitive Functions

Feuerstein’ approach works at a metacognitive level. It is grounded in a cognitive map which identifies the cognitive functions in their input, elaboration and output phases (Dawes, 2006). For the correct execution of each of these phases, the individuals require to employ diverse mechanisms and mental processes, defined as cognitive processes. Therefore, these processes may be understood as the set of internal, organized, and coordinated mental actions through which an individual processes received information, builds answers and conclusions, and communicates them. The cognitive functions of each of the three stages, as defined by Feuerstein (Feuerstein et. al, 1980), are shown in Table 1.

The input functions refer to the amount and quality of the data gathered by the individual before facing the solving process. After, the elaboration functions concern with the resolution process itself. Finally, the output functions relate with the correct and precise communication of the constructed solution.

During the first stage (i.e., input) the individual understands the problem, determines the ways to address it, and gathers relevant data, storing it in an organized and structured form. In this stage, the employed cognitive functions relate to perception, comparative conduct, language use, analysis of different information sources, use of spatial-temporal relationships, and impulsivity restraint.

Throughout the second stage (i.e., elaboration) the individual solves the problem by constructing an appropriate answer. For this, he analyzes and relates the previously acquired information, develops abstract representations, proposes hypothesis, validates them, and justifies his solution. Depending on the problem addressed, the solution may be not unique, what favors divergent reasoning.

In the last stage (i.e., output) the individual communicates the constructed solution. In this stage the individual must use clear and precise language in order to express the answer in a fast and correct way. In this stage the individual is expected to reason and consider before answering, overcoming random and trial-and-error type of responses.

Intervention Model

Feuerstein’s Instrumental Enrichment Program (FIE) focuses on the cognitive functions shown in Table 1. Supported by a battery of specifically designed instruments, its aim is to modify the cognitive structure of individuals, correcting deficient cognitive functions and improving the development of their operational abilities. The intervention we propose, while not based on FIE, it’s based on its underlying theoretical base.

The intervention model we propose is based on three main aspects. First, it exploits peer mediation through face-to-face collaboration (Zurita & Nussbaum, 2004; Zurita & Nussbaum, 2007), being the activities designed to be solved not individually, but collaboratively by a small group of students; second, it exploits the real world in which the students are immersed; and third, it combines this real world dimension with an abstract dimension. The underlying technological system that supports these requirements will be described in the next subsection. Following, we will describe the pedagogic model of the interventional activities.

Technological Model

The technological system used consists of an autonomous mobile robot and a set of wirelessly interconnected handheld devices; at the beginning of the activities, each student is supplied with one of these handhelds (Figure 1). As seen in the figure, the students and the robot share a common physical space.

This system allows exploiting the real dimension by means of a mobile robot able to interact with the physical world in two ways, sensing it and moving through it. Because mobility is fundamental for real-world interactions, the robot must be supplied with accurate navigation systems to allow it to precisely fulfill the educational activities. In relation to the students, the inherent mobility of the wirelessly interconnected handheld devices allows them to explore, measure, and analyze the physical environment.
The abstract dimension is provided by the handheld devices, where the students are requested to solve the problems. In this way, while a part of the problem is displayed physically, students model this reality to be able to provide abstract answers based on relevant physical information. Therefore, the handhelds graphical interfaces are the main mean through which each student interacts with the technological system.

Regarding face-to-face communication, for it to happen effectively among the students they must be allowed to keep a constant visual contact among them. In this sense, because of the reduced size and mobility of the wirelessly interconnected handheld devices, these become transparent tools during the activities, providing neither physical nor visual barriers among the students, allowing constant visual contact and face-to-face interactions to occur in a natural way.

<table>
<thead>
<tr>
<th>Technological Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction with students</td>
<td>The system must provide an interface to allow its interaction with each student.</td>
</tr>
<tr>
<td>Communication among devices</td>
<td>Seamless communication must be provided among all system’s devices, which is accomplished through wireless communication.</td>
</tr>
<tr>
<td>Interaction with real world</td>
<td>The system must be able to interact with the physical world, perceiving it and acting on it.</td>
</tr>
<tr>
<td>Autonomous Mobile Device</td>
<td>The system must provide a mobile device capable of traveling precisely and autonomously in its surrounding environment.</td>
</tr>
<tr>
<td>Face to face interaction</td>
<td>The system must not obstruct visual contact between the students. It must also provide the students of a shared physical space in which they can undertake effective face to face interactions, even when interacting with the technological system.</td>
</tr>
<tr>
<td>Student mobility</td>
<td>Students must be able to freely move around their surroundings to explore, measure, or analyze the interaction between the system and the environment, as well as to favor student’s face to face interaction.</td>
</tr>
</tbody>
</table>

*Table 2. Requirements of the technological system*

Finally, collaboration is supported based on the wireless interconnection among the handheld devices and the robot. Based on it, the systems is able to determine tasks, roles, and interaction rules among the students. Therefore, the system is capable of guiding collaboration, for example, by forcing every student to agree with group’s answer.
before continuing, or by distributing complementing roles among the students, generating a need in them to work together.

The summary of the previously addressed system requisites is shown in Table 2.

**Pedagogical Model**

The pedagogical model proposed is shown in Figure 2. First, the system presents the problem to be solved by the students, distributing relevant information among a set of handheld devices; each handheld device is assigned to a specific student. After the question has been proposed, the robot physically executes a set of movements that complement the problem’s information. After the robotic motion ends, the students are required to construct a common answer in their handheld devices. This answer must be constructed in a collaborative way, either by constructing an only answer in a shared and common process, or else by allowing each student to propose an individual answer, after what the group selects the one that better addresses the proposed problem.

Once a common answer has been constructed the systems prompts the students to confirm their consent with the constructed answer. In case agreement is not met, the students are impelled to propose a new set of answers, or else to collaboratively construct a new one. Once consensus has been achieved, the system evaluates the answer. According to its correctness, the system is able to provide feedback by displaying the correct answer in two forms, abstractly through the handheld devices, and physically by means of the robot. Finally (not shown in the figure), a theoretical reinforcement question is presented to the group. By means of it, the system helps the students to transfer, generalize, and abstract the concrete concepts just learned.

Regarding cognitive aspects, the real dimension, which takes advantage of the physical-world, requires the students to immerse, measure, and explore their surroundings in order to gather relevant information. This dimension fosters functions related to the input mental stage such as: **Clear Perception** and **Precise information gathering** since the students must distinguish relevant information from irrelevant stimuli and/or sensorial noise; **Systematic exploration of a learning situation** since they are required to extract and store appropriate parameters out from the relevant stimuli and information; and **Spatial organization** because of the inherent spatiality of the real-world dimension of the model.

Complementary, the abstract dimension and face-to-face collaboration stimulate functions relating to the elaboration and output mental stages. The abstract dimension takes advantage of graphical interfaces. Supported by these, it requires the students to transform information obtained from the physical-world into abstract models of reality, and/or to solve the physically shown problems in a virtual environment. This abstract dimension favors the development of **Internal mental representations** based on the need to represent reality using symbols and rules, as well as **Visual transport** when required to shift from the abstract to the real dimension and vice-versa.

Face-to-face collaboration relates mainly with output cognitive functions. Based on a collaborative problem-solving approach, face-to-face interactions give the students the chance to explain and verbalize their ideas, obtaining a
clearer perspective of the learned subjects (Gillies, 2006). Moreover, since the constructed answer must be approved by every group member, students are encouraged to propose and defend their ideas while asking their peers to explain and justify whatever ideas that they do not consent to or understand. In this respect, it has been shown that students working collaboratively increase their participation in group discussions, develop a higher level of discourse, interrupt less, and provide more valuable contributions to discussions (Shachar & Sharan, 1994). Hence, face-to-face collaboration promotes the development of Explicit communication, Uninhibited answer communication, and Answer Control. Furthermore, since the resolution is collaborative, students are encouraged to take roles and divide the problems into different complementing tasks, behaviors directly related to the Conduct planning cognitive function.

Finally, while the purpose and structure of any particular activity may vary according to the specific academic subjects addressed, in order to take advantage of the pedagogical model proposed, every activity must always maintain the collaborative scheme as well as the physical-abstract dimensional requirement.

Implementation

Based on the proposed intervention model, we situated our study in high school physics, specifically in graph plotting. We designed an activity named Graph Plotter whose academic purpose is to develop and enhance graphical skills among the students, such as constructing kinematics graphs of velocity and distance, and its reading and interpretation. To achieve this, the robot executes a set of movements to be plotted by the students. This activity, along with the cognitive processes addressed by it, is detailed in the next subsections.

Graph Plotter Activity

The educational activity consists of a group of three students that have to graph different linear movements performed by a mobile robot in a collaborative way. For this, each student is given a WiFi enabled PDA which establishes a wireless network with the robot and with the rest of the PDA devices. The tools provided to the students are a set of on-screen blank virtual axes and a set of on-screen synchronized chronometers, along with a measuring tape for them to measure real distances.

The motions of the robot are one-dimensional (i.e., the robot moves along through a straight line), varying its speed, acceleration, and direction. The types of graphs which the students are requested to plot are either position versus time or velocity versus time. On group demand, the robot can repeat the performed motion for the students to be able to gather all the required data and validate their plots.

To build the graphs the students are given measuring tapes to calculate, as precisely as possible, the relevant distances traveled by the robot. In addition, an on-screen chronometer, synchronized with the starting of the robot’s movement is displayed which allows them to determine the time of occurrence of each of the different movement events. Concerning the virtual dimension, where the solution must be developed, along with the chronometer a synchronous color sweep graphically shows the movement’s current time (Figure 4b).

Figure 3 shows a diagram of the activity flow. Upon starting the activity, and after the exercise’s objective has been explained (Figure 4a), the robot performs a set of linear movements on the floor (Figure 5) which the students have to plot using the stylus on their PDA screens (Figure 4b–4c). The group may ask the robot to repeat the performed motion as many times as necessary. When all the group members finish their plot, the group is presented every individual member’s solution, so as for the group to select their final collective answer (Figure 4d), what favors discussion, consideration and reasoning (Zurita & Nussbaum, 2004); if different answers are chosen, the system requires them to come to an agreement (Cortez et al. 2005). Independent of the number of times the students fail to achieve a consensus, the system will keep demanding consent before continuing with the activity.

When the collective answer has been chosen, the system autonomously determines whether the proposed answer was correct or not, informing the students. How an answer is evaluated will be described in the next paragraphs. In case the answer was correct, the system shows in each PDA screen the exact solution superposed over the students’ answer (Figure 4e); otherwise the system displays the students’ plot alone on the PDA screens. After this, the robot
repeats once again the movement, while the system synchronously renders the correct graph on each student’s PDA (Figure 4f). This allows the students to observe the relationship between the movement in the physical dimension and its abstract representation on the PDA’s screen.

![Diagram of Graph Plotter Activity flow](Figure 3. Diagram of Graph Plotter Activity flow (Mitnik et al. 2009))

Finally, independent of whether the graph was correct or wrong, a reinforcement question is asked to the students for them to infer knowledge from the previous experience. While in the plotting stage the students are asked to transform

Figure 4. Screenshots of Graph Plotter. a) Exercise goal. b) Graph plotting while the robot moves. c) Graph editing. d) Selection of the group answer. e) Exact answer superposed to the correct group answer. f) Correct answer is rendered in each PDA. g) Symbolic-Textual reinforcement question. h) Symbolic-Symbolic reinforcement question.
reality into an abstract model, in this final stage the plot is used to extract and analyze information. All reinforcement questions are multiple choice questions which refer to the last motion performed by the robot, addressing issues such as: displacement in a time interval, instant and average velocities, instant accelerations, textual descriptions of the motion (Figure 4g), maximum/minimum variable values, and transformation into other kinematics graph (Figure 4h), among others. This is how, while the question has always a symbolic component (i.e., the graph), the required answer may be either textual (Figure 4g), numerical or symbolical (Figure 4h). Once again, for the group to respond every individual member must agree with this collective answer.

The whole activity ends after the students have answered all the exercises planned for that activity session. These previously planned exercises are supplied to the system using an exercise file.

Regarding technological issues, all the software used in the activity was designed and developed specifically for it, including user interfaces, inter-device synchronization, collaboration schemes, and the activity logic. The physical robot used was a commercially available one.

In order to evaluate whether the proposed graph is correct, the system compares it with the exact solution performing a point-by-point error calculation. Since plotting a perfect graph is extremely difficult, the system determines that a proposed graph is correct if the sum of the point-by-point error falls below a predefined threshold.

In relation to inter-device clock-synchronization, inaccurate time synchronization between the PDAs and the robot would translate into erroneous or delayed student measurements. Hence, an accurate clock-synchronization algorithm was developed to overcome the variable latency-time and communication delays inherent to wireless network connections. Details of the algorithms used may be found in (Mitnik et al. 2008).

**Graph Plotter Experimental Design**

This activity was tested with 10th grade students that had been taught kinematics during the same year. The activity was used by 24 16-year-old students. Throughout the two weeks of the experiment, the students completed five working sessions of 60 minutes each. Each session consisted of a set of at most seven difficulty-increasing exercises. Throughout the activity sessions the exercises continually increased in complexity. At the beginning of each session students were randomly assigned into groups of three members. Since many of these students came from different classrooms, many of them had never interacted before with their group mates. All through the activity, video recording and in-site observations were performed. The data collected by these two methods was analyzed based on the cognitive functions described in Table 1.

During each session, exercises of position and velocity graphs were combined. Teacher assistance was provided throughout the whole experiment, yet only when required by the students themselves. Other than this, during the study the students received no other form of kinematics instruction.

Since academic learning is an important issue in the intervention, accurately determining the learning accomplished by the students is prior. In order to assess this learning, a pretest-posttest design was used, administering the pretest the day before the intervention began, and administering the posttest the day after the intervention concluded. The instrument used in both tests was the *Test of Understanding Graphs in Kinematics* (TUG-K) (Beichner, 1994), a 21 multiple choice test designed specifically to measure graph interpretation skills; its reliability and internal consistency have been validated (Beichner, 1994). A relevant issue of this test is the fact that it focuses on the seven more difficult aspects of graph understanding, aspects commonly not understood by the students after traditional kinematics classes. Hence, this test is a very difficult one and scores are expected to be low.

Finally, motivation was also measured based on the results of a post activity motivational survey. Each of the survey’s questions had five possible answers: --, -, 0, +, and ++; “--” meaning “very little” while “++” meant “very much”. The questions of the post-activity survey were the following:

- Did you like this activity?
- How motivated did you felt during the activity?
- How entertaining was the activity?
- How challenging was the activity?
Results

After the intervention ended, we observed that the students spent 10 to 15 minutes solving each exercise. This solving time was consistent through all the intervention sessions even though the exercises continually increased in complexity. As mentioned before, learning outcomes were tested based on a specific graph understanding test. Cognitive aspects, on the other hand, were analyzed based on qualitative in-site and video recording observations.

Learning results

To determine how this activity achieved learning and graph understanding throughout the students, a pretest and a posttest were administered. Even though the activities were designed to be solved in groups, these tests were answered individually by each student in order to identify the independent learning outcomes. The average pretest score was of 5.11 correct answers. This mean score increased after the intervention to a score of 7.11 correct answers. By performing paired samples T-test we could determine that this increment of two correct answers, corresponding to a 39% increase, was statistically significant (p<=0.01), presenting a medium effect size (Cohen’s d = 0.785) (Mitnik et al. 2009).

Motivational results

Regarding motivation, students were observed to be motivated when participating in the activities and during all the resolution process. In relation to the elements that promoted motivation, it was observed that motivation was supported by the students’ need and possibility to immerse into the activity. Supported by this immersion, the students were not just spectators, but became relevant actors of the activity, developing a great commitment to their team and toward the activity resolution process. This motivation was seen to prolong all throughout the intervention. This is how during the last session many of the students expressed their wish to continue working with this kind of activities.

These observations were consistent with the results of the post activity survey. As mentioned before, each of the four questions had five possible answers: --, -, 0, +, and ++. In order to quantify these results we assigned these answers the scores -2, -1, 0, 1, and 2. The mean score of each of the survey questions is shown in Table 3. As seen, the global mean value was found to be 1.37, signaling that in average the students found the activity almost amid “motivating” and “very motivating”.

<table>
<thead>
<tr>
<th>Question</th>
<th>Average value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you like this activity?</td>
<td>1.37</td>
</tr>
<tr>
<td>How motivated did you feel during the activity?</td>
<td>1.29</td>
</tr>
<tr>
<td>How entertaining was the activity?</td>
<td>1.43</td>
</tr>
<tr>
<td>How challenging was the activity?</td>
<td>1.19</td>
</tr>
<tr>
<td>All four questions</td>
<td>1.37</td>
</tr>
</tbody>
</table>

Table 3. Mean values of the motivational survey. The questions had five possible answers: -2, -1, 0, 1, and 2

Graph Plotter Observed Cognitive Processes

From the qualitative observations obtained for the previously described activity, and following the cognitive functions of each of the three stages, Input, Elaboration and Output, as defined by Feuerstein (Feuerstein et. al, 1980) we reached to the following results.

Input

Clear perception. The failure of this cognitive process would be evidenced by an incomplete or erroneous information gathering. Since the activity and the environment were both, novel and complex for the students, they
were forced to learn how to observe and extract relevant data from the problem. Even though most students were unable to determine useful information on the first session, all of them were seen to develop this observance ability throughout the study. Students learned to differentiate relevant from irrelevant data. For example, after the first session students started focusing on the instants when speed or direction changed, rather than on the intervals during which the robot moved in a steady way. Based on this, they learned that in order to gather the information required in a correct and complete way, they did not needed to be measuring distances constantly, but required just a few specific measurements.

**Precise information gathering.** Initially, the groups presented extremely uncoordinated processes when gathering information. Even when the whole group knew what information was needed, lack of coordination appeared, shown when two students gathered the same information, also leaving some pieces of information ungathered. This forced the students to observe the robotic motion many times before being able to solve the problem. Nevertheless, throughout the study students were seen to develop collaborative strategies which allowed them to gather information in an organized and efficient way. This is how the students assigned each group member a particular role in the gathering process, showing not only collaboration, but teamwork. In a particular group strategy, for example, one student notified whenever the robot changed speed, another one marked the spot where this happened, and the last one recorded the time of this occurrence.

**Systematic exploration of a learning situation.** While the strategies developed proved to be effective, these would have been useless unless the students developed a way to organize and store the gathered information. Since relevant information consisted of two parts, the time and the position of the robot when its speed changed, students were seen to store information as pairs of data. Nevertheless, because of the strategies used, one student possessed the timing data while another stored the corresponding distances. Hence, each student stored his/her information in chronological order to pair it with his/her teammate’s data.

**Spatial organization.** Students learned how to position themselves in order to fulfill each role requirement in their collaborative work. In the first sessions the physical arrangement of the groups offered little help concerning the data gathering process (Figure 5a), nevertheless this arrangement evolved to the one seen in Figure 5b, where each student is placed according to his/her role. For example in Figure 5b the rightmost student is placed far from her group to observe when the robot executes its speed changes.

![Figure 5](image_url)

*Figure 5. Students working with Graph Plotter Activity. a) During first sessions, the physical arrangement of the group offers little help to the data gathering process. b) Physical arrangement after strategies were developed and roles were assigned. The rightmost girl is placed where the robot changes its speed. The middle girl focuses on her PDA’s virtual chronometer, while the leftmost girl, in charge of notifying speed changes, maintains her attention on the robot.*

**Temporal orientation.** In combination with Spatial Organization, this ability was fostered based on the nature of the problems, where it was essential to relate events such as speed changes, distances, and times. For example, students understood that velocity depends not only on distance or time, but on the relation between these.
Object conservation. This ability is supported by the invariant attributes of the kinematical concepts. For example, students could observe that the distance traveled is independent of the velocity at which it was traveled, or the time spent traveling it.

Organization of information. The use of different information sources is attained twofold. On the one hand, students obtained information presented by the robot itself, measuring it by means of measuring tapes and a chronometer. On the other hand, each teammate acted as a new information source, not only because of his/her adopted role, but because of his personal thoughts and ideas regarding the resolution of the problem.

Elaboration

Perception and definition of a problem. Throughout the activity, the types of requested graphs were randomly assigned to be either position or velocity vs. time. Therefore, students learned to discriminate what type of graph was being asked, what information was required, and how was this data represented in each case.

Selection of relevant knowledge. Position and velocity graphs have different meanings regarding slope, area under the curve, and height. After the study, we observed that the students became able to correctly analyze graphs which represented very different movements, even when their plots looked very similar. This was accomplished because the students developed the ability to switch between the two kinematics models learned, according to the proposed problem.

Internal mental representation. Expanding on the previous paragraph, because kinematic models and graphic representations are abstract representations of reality, the correct use of internal mental representations was crucial for the correct resolution of the activity. Without the ability to use abstract symbols and concepts to represent reality, the problems simply could not have been solved.

Perceptual organization and structure. During the activities, independent of the type of graph, the scale of the axis changed from problem to problem. Therefore, the students learned not only to plot their data, but also how to scale and represent it according to the set of axis posed in the problem.

Conduct planning. The collaborative strategies developed by the students, mentioned in the previous section, focused not only on the data gathering process, but on the resolution process as a whole. It was seen that, initially, the students watched the robot’s movements in order to identify the positions and number of speed changes. After this they distributed roles among them, each student placing himself at his corresponding location (Figure 5b). Then they asked the robot to repeat the movement. During this first repetition the students collected and measured relevant data. Once this was accomplished, they gathered together to assemble every member’s data and plot their own graphs. In case one or more members were uncertain of the plotted answer, they asked the robot to perform again, time when the students checked the correctness of their response.

Hypothetical thinking and Logic evidence. The development of these cognitive functions was observed throughout the discussions of the students, time when each member presented his hypotheses or ways of thinking. Based on the fact that the students were able to observe as many times as necessary the robotic movement, they were able to validate or reject the proposed graphs. Moreover, since every group member was required to agree with the group solution, the students developed the ability to justify their solutions in order to convince the other members of the correctness of the defended graph.

Output

The output-stage cognitive functions are mostly developed due to the collaborative structure of the activity. Because consensus was demanded, plenty of dialogue aroused among the students.

Explicit communication. This cognitive function was fostered each time a group member needed to defend or present his ideas to a teammate. In these common discussions it was seen how the students learned to adapt their explanations according to their listener needs. An example of these were dialogues where one student explained to
another using knowledge previously acquired by both of them, using phrases such as “remember when the robot moved slowly, the plot was like this... well, now it also moved slowly, so the plot must be similar”.

Uninhibited answer communication. Initially, the students were quite shy when the time came to explain their points of view. However, as sessions passed, students were seen to become more confident on their reasoning. Based on the constant dialogues, the students developed the ability to construct better and more precise explanations.

Answer control. Random and trial-and-error answers, as well as student impulsiveness, were completely inhibited thanks to the consensus requirement. Since each group member was required to agree with the proposed solution, and each member was seen to demand explanations before agreeing on an answer, each student was forced to reason and fundament their answers before proposing them. This reasoned proposals resulted in the construction of better explanations and faster consensus reaching.

Accurate and precise answer. Finally, this cognitive function was seen to develop by means of the answers themselves. When comparing the answers of a student in the first, middle and final sessions, we could observe that the exactitude of the plots improved significantly.

Conclusions

The technologically-based pedagogical model presented opens new perspectives regarding the use and potential of computer assisted instruction. In particular, this novel activity model has shown to accomplish a dual academic purpose: it allows the students to correct and enhance their cognitive processes while also addressing school curriculum topics, such as physics in the presented case.

From the qualitative analysis we can conclude that in order for the students to achieve cognitive development throughout all the three stages of the mental process, each of the constructs of our model is required, these are: a real dimension, a virtual dimension, and a collaborative design.

First, the real dimension provided by the physical robot focused mainly on the development of the input cognitive functions. Based on this real-world exploring capability the students were able to train their abilities to precisely gather information from their environment, to differentiate relevant from irrelevant data, to develop strategies, to coordinate the group members all through this gathering process, and to organize and store this collected information in a useful manner. All these are examples of clear perception, precise information gathering, systematic exploration of a learning situation, and organization of information, among others. This cognitive training showed visible as the students progressed from an uncoordinated data gathering process in which relevant data was commonly lost, to a distributed role scheme were each student contributed with a complementing part of the required data.

Second, the virtual dimension provided by the PDAs focused mainly on the development of the elaboration cognitive functions; this virtual dimension provided the activity with the ability to exploit the use of real-abstract relations. By requiring the students to develop mental representations of their physical reality, to model it using abstract symbols and concepts, and to determine the appropriate model to use in each problem (e.g., either position or velocity representations), students trained cognitive functions such as selection of relevant knowledge, internal mental representation, perception and definition of a problem, and perceptual organization and structure. On the other hand, based on the collaborative nature of the activity, its pedagogical design, and because the robot can act as an endless tutor, students were given the opportunity to propose as many hypothesis and ways of thinking as they want, being able to validate or reject them through the repeated observation of the robot’s movements. These allowed the students to develop conduct planning, hypothetical thinking, and logic evidence.

Finally, the face-to-face collaborative design of the activity promoted the development of the output cognitive functions. Because consensus was demanded, plenty of dialogue aroused among the students. Students learned not only to verbalize their ideas, but also to adapt their explanations according to their listeners’ knowledge and needs. Since each student needed to defend and logically validate his answer in front of his teammates, students were encouraged to reason before delivering an answer; this allowed impulsivity to be inhibited while insight among the students was intensified. All these are examples of how explicit communication, uninhibited answer communication, answer control, and accurate and precise answer were enhanced.
Regarding the cognitive enhancement as a whole, this approach differs from Feuerstein’s Instrumental Enrichment Program (FIE) in two ways. On the one hand, while FIE’s sole purpose is to correct cognitive deficiencies, not addressing any curriculum objectives, the intervention model here presented allows the teaching of specific academic issues, as for example, kinematics. Because cognitive modification is accomplished in parallel with curriculum learning, this cognitive training needs no extra classroom time, making it feasible to be inserted into the school curriculum. On the other hand, FIE requires the child to consciously dedicate time and effort in the program. Our intervention model, on the contrary, allows cognitive training to be unnoticed by the students since they are focused on solving the activities rather than on cognition training. Hence, our model achieves cognitive modification as a natural result of the activity.

Concerning curriculum issues, graph understanding showed a significant increase after the intervention. Since the measuring instrument was a rather difficult test, the 39% increase in the students’ scores sustains the claim that not only cognitive modification was accomplished, but also curriculum learning.

Finally, students were seen to learn in a rather stimulating environment, as confirmed by their will to continue participating in this kind of activities. This finding was also supported by the results of the motivational survey which showed that in average the students found the activity to be either “motivating” or “very motivating”. It is important to say that, while motivation may be attributed to the newness of the technologies used, we believe this to be relevant, but not a key issue. Based on the observations made, and on the fact that motivation remained constant all-through the intervention, we believe that the key issue that promoted motivation among the students was the fact that the students’ needed to immerse into the activity setting, becoming relevant actors of the activity, and developing a great commitment toward their team and toward the activity resolution process.

To conclude, regarding the future work of this research, it focuses on the design and development of new educational activities which take advantage of the pedagogic and technologic model here proposed. Examples of activity subjects can be found in the field of mathematics, where the robot may be used to illustrate concepts of geometry, vectorial calculus, and path planning, among others. In relation to the measurement of cognitive functions, part of the future work focuses on the development or modification of measuring instruments in order to be able to obtain quantitative measures of the cognitive progress of the students.

References


Are the Functions of Teachers in e-Learning and Face-to-Face Learning Environments Really Different?

Laura Alonso Díaz* and Florentino Blázquez Entonado

Department of Education Sciences, University of Extremadura, Av. Universidad s/n, 10071, Cáceres, Spain

laulonso@unex.es // blazento@unex.es

*Corresponding author

ABSTRACT

The main purpose of this study is not to compare online and traditional face-to-face instruction merely to prove which one is better, but rather it aims to highlight some of the possible risks and strengths which may help to improve the role of teachers in both methods. The scene consisted of various thematic blocks from a training programme, with teachers who taught two different groups of students, one of them face-to-face and the other online. The study was designed using a quantitative and qualitative methodological combination, and focuses on the dimensions of “theoretical content”, “practical content”, “tutor/student interaction” and “design” of the training activity. As a general conclusion, no important differences were observed in the functions of the teacher in the two teaching methods, face-to-face and online; any differences that might exist were usually a consequence of teacher involvement and of the commitment of the institution in programming the learning process.

Keywords

E-learning, Face-to-face learning, Computer-mediated communication, Online teacher, Traditional teacher

Introduction

Although e-learning is still in its infancy, the knowledge acquired by teachers who use online and face-to-face methods can be of great use in improving both types of teaching, which is the reason why researchers nowadays study issues related to these teaching methods, e.g. Urtel (2008) and Georgouli et al. (2008). It is not simply a question of retaining traditional teaching methods such as the master class and applying e-learning techniques to gain access to more information. Nor does it mean involving the students in the same learning methodology using a different medium.

Considerable progress must still be made in order to enable today’s society to take full advantage of the potential of online learning. Several researchers (Wilcox & Wojnar, 2000; Mason, 2003; Rovai, 2004; Salmon, 2004; Kearsley, 2005; Cabero, 2006; García Aretio et al., 2006, among others) have reported on the peculiarities in design, contents, activities, interaction, tools and evaluation processes in face-to-face and online modes of teaching. On comparing the two methodologies, people may mistakenly regard the two processes as similar when in fact they should be seen as different from the outset. However, it is useful to carry out comparative research which might lead to improvements in each type of learning model. Coates et al. (2004) have pointed out that it is negative to explain only the differences between face-to-face and online methods and not the basic attitudes which form the starting point for each model.

At the beginning of the 1990’s education by correspondence was criticized on the grounds that it reduced education to a mere process of industrial production (Peters, 1993). Shaw (2001), however, believes that problems only arise when conventional teaching methods are simply adapted to distance learning; indeed, Johnson et al. (2000) found that there are no significant differences between the two methods when success factors are determined. In this way, students can learn just as effectively in either of the two formats, whatever their style of learning, providing that the teaching is conducted with an adult learning theory and has instructional design guides. In this respect, online learning is especially useful for students who, for reasons of work, family or social commitments, require a different way of learning (Jeffcoat & Golek, 2004).

Blankson and Kyei-Blankson (2008), among others, have investigated the extent to which students are satisfied with online, face-to-face and blended courses. They decided to integrate synchronous online discussions with traditional face-to-face instruction, and the results of their study suggested that students were generally satisfied with the blended course format. However, Lim et al. (2008) conducted an empirical investigation of student achievement and satisfaction in different learning styles, and found that no significant differences existed between online learning and traditional learning groups. So and Brush (2008), having also studied student perceptions of satisfaction in a blended...
learning environment, discovered that students who perceived high levels of collaborative learning tended to be more satisfied with their distance course than those who perceived low levels of collaborative learning.

Reisetter et al. (2007) examined whether online learners and face-to-face learners are equally satisfied with the quality of their learning; their findings showed that both learning styles scored equally with regard to learning outcomes and satisfaction, despite the fact that each style has decidedly different learning experiences. This study offers insight into the nature of the experience of online learning, and suggests that online course designers focus their attention on particular elements that support the unique experiences of students who select this learning method.

The efficacy of face-to-face and online learning was compared by Solimeno et al. (2008). Overall, their results showed that asynchronous collaborative learning online can increase professional competences normally acquired only in small face-to-face educational settings; they report that online learning can be used to provide innovative educational opportunities to fit the particular needs of students who have time management problems in their learning strategies, with low anxiety, high problem solving efficacy.

A field experiment carried out by Hui et al. (2008) compares the effectiveness and satisfaction associated with technology-assisted learning with that of face-to-face learning. It showed that technology-assisted learning improves students’ acquisition of the kind of knowledge which requires abstract conceptualization and reflective observation, but adversely affects students’ ability to obtain knowledge which requires concrete experience. Technology-assisted learning is better for vocabulary learning than face-to-face learning, but it is comparatively less effective in developing listening comprehension skills.

It is clear that to compare the advantages and disadvantages of each system would require an examination of the needs of a diversified population. Some researchers have concluded, in this respect, that it is necessary to design flexible courses that integrate techniques from both face-to-face and online methods (Delfino & Persico, 2007). Wuensch et al. (2008) evaluated the pedagogical characteristics of their most recently completed face-to-face class and their most recently completed online class. The results showed that students rate online classes as far superior to face-to-face classes in terms of convenience and in permitting self-pacing, but they also rate online classes as inferior in a number of other ways. Online and face-to-face instructional formats, then, each have their own strengths and weaknesses. The authors cited have detailed these strengths and weaknesses with the aim of improving both methods of teaching by reducing the weaknesses and maintaining the strengths. These studies confirm the relevance of what is known as “blended learning”, which consists of the combination of face-to-face and distance teaching/learning methodologies. According to Berger et al. (2008), online and face-to-face environments play different and complementary roles. In this way, the development of “blended learning” as a grounding area will enable teachers to design, develop and deliver effective mixed programmes (Chew, 2008). However, Jackson and Helms (2008) have found that hybrid classes continued to exhibit the same weaknesses of the online format, and that the addition of face-to-face interaction does not minimize weaknesses.

The University of Extremadura organizes courses for obtaining the Certificate in Pedagogical Aptitude for those graduates who wish to become secondary school teachers. These courses include a phase of general psychopedagogical training, in face-to-face or online mode, and it is in this phase that our study has been conducted. Online teaching is offered with the aim, on the one hand, of providing an alternative for those students who have difficulties in attending classes in the traditional manner, and on the other, of promoting teaching methods that allow teachers to make use of new advances in communication technology. The teaching process is supported by multimedia didactic material which the student must study with a computer, via Internet, using the University’s Moodle platform. This material has the same psycho-pedagogical contents as that used in the face-to-face classes. The duration is the same as for class attendance, based on the estimate that a student will need to invest a maximum of 100 hours of work and participation in the various activities required for a pass in this phase. The students are also offered at least two non-compulsory though recommendable face-to-face sessions. For each of the study topics, the student has the support of a designated teacher who is responsible for clarifying doubts and questions related to the course contents. Communication between teachers and students, and students and students, is made by email, chats, forums and telephone, depending on the purpose. The teaching group which has been in charge of this programme for a number of years, and of which the researchers in this study form part, is a community of practice with a shared desire to introduce innovations in online teaching. Learning to teach is a process built on understanding, developing and effectively using personal characteristics in relations with the students.
This secondary school teacher training programme was studied in the year 2007. The scene set for our research consisted of three different thematic blocks or subjects of the training programme of the course that were taught by the same three teachers to two different groups of students, the face-to-face students and the online students (Figure 1).

The purpose of this study is not to compare online and face-to-face instruction merely to prove which one is better, but rather it aims to determine possible weaknesses and strengths of each of the two methods with the aim of improving the role of teachers in online and face-to-face learning. The main objective of this research, therefore, was to determine whether differences exist in work-related tasks carried out by teachers of online and face-to-face systems, with regard to:

- **Theoretical Content**: conceptual theoretical content of the training action and its online structure;
- **Practical Content or Activities**: practical activities useful for understanding the theoretical content;
- **Interaction**: the process of the relation which occurs between teachers and students throughout e-training, and its implications;
- **Design**: management and administration aspects related to the distribution of content with regard to space and time, instructions, technical problems etc.

The following hypotheses were established:
1. Satisfaction and efficacy in the study of *theoretical content, practical content*, and teaching *design* is higher in the online training programme than in the face-to-face training programme.
2. Satisfaction and efficacy in *interaction* processes is greater in the face-to-face training programme as compared to the online training programme.

**Method**

A combination of both quantitative and qualitative methods was used.

Quantitative methodology was considered pertinent from the point of view of comparing concrete aspects of the same teaching programme taught in two different modes, i.e., face-to-face and online. It was not only effective in establishing comparisons between the two modes, but also facilitated the treatment of data on large populations.
To ensure greater rigour in the methodology and reliability of the data obtained, qualitative data were added to the statistical results. In this way the two methods complemented each other to enhance the quality of the study.

A total of 255 subjects participated in the study, two of which are renowned experts in online and distance education: Dr. Julio Cabero Almenara (Professor of Didactics and School Organization at the University of Seville) and Dr. Lorenzo García Aretio (Chairman of UNESCO Distance Education in Spain). The number of students participating in the training programme in its online and face-to-face versions was 250 (129 online and 121 face-to-face), plus three teachers who took part in both types of teaching.

For the collection of data, a triangulated technique of closed questionnaires, semi-structured interviews, and discussion groups was used, as shown in Figure 2.

![Figure 2. Techniques](image)

The process of collection and analysis of data, and interactivity among the various research tasks, is presented in Figure 3.

![Figure 3. Process of collection and analysis of data](image)

**Closed questionnaire**

A closed questionnaire was prepared to gather data on the following variables:

- A nominal variable: “group”; with two values: “face-to-face group” and “online group”.

334
Twenty quantitative variables grouped in four dimensions: “theoretical content” of the subject-matter; “practical content” of the subject-matter; “interaction” with teachers; “design” of the training programme.

The initial questions on the questionnaire were prepared after a review of the literature and of previous studies (Wilcox & Wojnar, 2000; Mason, 2003; Rovai, 2004; Salmon, 2004; Kearsley, 2005; Cabero, 2006; García Aretio et al., 2006, Blázquez & Alonso, 2006). The first version of the questionnaire, with 26 questions, was submitted for judgment by the experts, by email and ordinary mail. The experts evaluated whether the 26 questions suggested were more or less pertinent using a Likert-type scale with five different values (1 = not pertinent… 5 = very pertinent); they also filled in a section suggesting possible modifications for each question. Finally, the questions that were considered by the experts to have a value lower than 4 were removed from the questionnaire.

The final questionnaire consisted of 20 questions, of which 18 solicited information on levels of satisfaction and efficacy of “theoretical content” (items 1 to 6), “activities” (items 7 to 12), and “tutor/student interaction” (items 13 to 18) for each of the three thematic blocks. Two questions (items 19-20) addressed levels of satisfaction and efficacy of the “design” of the training programme in each of the online and face-to-face teaching modes (Table 1).

<table>
<thead>
<tr>
<th>Items</th>
<th>Content of the items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>Satisfaction with the theoretical content of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>4-6</td>
<td>Efficacy in the study of the theoretical content of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>7-9</td>
<td>Satisfaction with the practical content of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>10-12</td>
<td>Efficacy in the study of the practical content of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>13-15</td>
<td>Satisfaction with the interaction of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>16-18</td>
<td>Efficacy in the study of the interaction of thematic blocks 1, 2 and 3</td>
</tr>
<tr>
<td>19</td>
<td>Satisfaction with the design of the training programme</td>
</tr>
<tr>
<td>20</td>
<td>Efficacy in the study of the design of the training programme</td>
</tr>
</tbody>
</table>

With regard to the precision of the instrument designed, that is, its reliability, Cronbach’s α coefficient was used. By this means it was verified that the reliability of the questionnaire designed was α = 0.916, so it was assumed that reliability was high and that the questionnaire would measure the previously described theoretical dimensions with precision.

At the end of the training programme the 121 face-to-face students and the 129 online students completed the questionnaire.

**Discussion groups and semi-structured interviews**

The discussion groups and semi-structured interviews were designed to complement the quantitative data of the questionnaire. In this way, the questions were designed following the same variables used previously in the questionnaire, as they had been validated previously by the experts. The participants were asked about satisfaction and efficacy of the “theoretical content”, “practical content” and “tutor/student interaction” of the training activity in the three thematic blocks analyzed; and about the “design” of the online and face-to-face teaching activities.

In the preparation of the discussion groups and interviews, the “funnel” strategy was employed, which consists of questions of a more general nature and other more specific ones on each of the categories. This strategy was only used when it was necessary to delve more deeply to obtain information (Buendia et al. 1997). Two researchers moderated the discussion groups and carried out the interviews, which were recorded with the permission of the participants.

The following discussion groups and interviews were held:

1. First discussion group: eight students from the training programme (four face-to-face and four online) took part. The aim of this discussion group was to ascertain students’ levels of satisfaction with and efficacy of “theoretical content” in each of the two teaching modes. (E.g.: “Do you feel satisfied with the theoretical content of the first thematic block? Why?”)
2. Second discussion group: eight students from the training programme (four face-to-face and four online) and the three teachers who taught the two teaching modes took part. The aim of this group was for teachers and students to take part in a discussion on “theoretical content”, “activities” and “tutor/student interaction” in the three thematic blocks analyzed, and on the “organizational design” of the online and face-to-face training activities. (E.g.: “Do you think that there has been good interaction between you and the teachers in the course?”)

3. Semi-structured interview with teachers: teachers were questioned on their levels of satisfaction and efficacy with regard to the “theoretical content”, “activities” and “tutor/student interaction” in their face-to-face and online thematic block. They were also asked about their levels of satisfaction and efficacy regarding the “organizational design” of each of the teaching modes (online and face-to-face). (E.g.: “Do you think that the design you have prepared for the online course is as good as the design for the face-to-face course?”)

4. Semi-structured interview with experts: experts were asked in which respect they considered that the work functions of the face-to-face teacher differed from that of the e-learning teacher. (E.g.: “Do you think that there are differences in the way that face-to-face teacher and online teachers develop their tasks?”)

Results and discussion

Results of closed questionnaires

For the analysis of the variables with inferential statistics, the Mann-Whitney U-test was used, since it is the non-parametric test used for two independent groups. It was necessary to use non-parametric tests because in no case was the assumption of homocedasticity fulfilled, as was previously ascertained by Levene’s test.

The table below summarises only those items which were confirmed with 95% confidence, together with the result of the Mann-Whitney U-test and its significance (see Table 2).

<table>
<thead>
<tr>
<th>ITEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction with the theoretical content of thematic block 1</td>
</tr>
<tr>
<td>Efficacy in the study of the theoretical content of thematic block 1</td>
</tr>
<tr>
<td>Satisfaction with the practical content of thematic block 1</td>
</tr>
<tr>
<td>Satisfaction with the practical content of thematic block 2</td>
</tr>
<tr>
<td>Efficacy in the study of the practical content of thematic block 1</td>
</tr>
<tr>
<td>Efficacy in the study of the interaction in thematic block 2</td>
</tr>
<tr>
<td>Satisfaction with the design of the training programme</td>
</tr>
</tbody>
</table>

With regard to the items on satisfaction with the theoretical content, there were significant differences between online and face-to-face training only in the first thematic block, online structuring of theoretical content being higher (U= 5640.500 and p<0.05).

For efficacy in the study of the theoretical content, only design and structuring of the theory of thematic block one was significantly better in the online group than in the face-to-face group (U= 6414.500 and p<0.05).

Satisfaction with practical content in thematic block one was higher in the online version of the course (U=5809.000 and p<0.05). Similarly, practical content of the second thematic block was more satisfactory in online mode than face-to-face (U=6435.500 and p<0.05).

It was found that for efficacy in the study of practical content, thematic block one was more efficacious in online mode than in face-to-face mode (U=5767.000 and p<0.05).

With regard to efficacy in the study of the process of interaction with teachers, the results show that thematic block two was more efficacious in face-to-face mode than in online mode (U= 6244.000 and p<0.05).

Finally, it was found that the online students were significantly more satisfied with the design of the training programme than the face-to-face students (U= 6704 and p<0.05).
Results of discussion groups and semi-structured interviews

For the analysis of the discussion groups (D.G.) and the semi-structured interviews the qualitative method of “Content Analysis” was used, that is, a classic research technique in Social Sciences which is applied to non-structured or little-structured informative data, such as in our case.

In line with the procedure used by, amongst others, Miles and Huberman (1994) and Lacey and Luff (2001), the process of analysis of data began with the transcription of the recording of the interviews and the focus groups. With the aid of “AQUAD 6” software, a computer programme was created to serve as a base for theoretical approximation of a qualitative nature, in which the data were organized in easily recoverable sections. Then the initial codification phase was begun, which permitted us to ascertain inductively the principal categories of the study. Each of the principal categories was analyzed in order to find inferior categories or internal sub-categories. There were times when certain data that were not initially contemplated were seen as new categories, so that it was necessary to recode some levels which had been previously analyzed. Finally emergent topics and concepts were identified, which permitted us to recode and develop better defined categories, keeping in mind at all times the idea of a recurrent, flexible and iterative process.

As is shown below, firstly the interviews with teachers and the discussion group between teachers and students were analyzed; and secondly, the interviews with the experts were examined.

Results of the semi-structured interviews with teachers and the discussion group between teachers and students

This information was obtained from the three interviews with the teachers, the first discussion group between face-to-face students and online students, and the second discussion group between teachers and students in both teaching modes. In this way the second framework of categories and sub-categories was established. This is presented in Table 3, which also shows the number of comments or interventions made by face-to-face students, online students and teachers for each category and subcategory.

Table 3. Functions differentials between e-teachers and face-to-face teachers (views of teachers and students)

<table>
<thead>
<tr>
<th>CAT SUBCATEG.</th>
<th>DESCRIPTION</th>
<th>F2F Stdnt.</th>
<th>On-L. Stdnt.</th>
<th>Teacher</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTENT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous</td>
<td>Anticipated reflection on structure and motivating aspects</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>21</td>
<td>16%</td>
</tr>
<tr>
<td>approach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objectives</td>
<td>Aimed at comprehension, development of attitudes and orientation of the student</td>
<td>7</td>
<td>2</td>
<td>11</td>
<td>20</td>
<td>15%</td>
</tr>
<tr>
<td>Elements of</td>
<td>Main ideas, extended material, terminology, summary, tests, map, subject-matter objectives, references</td>
<td>14</td>
<td>10</td>
<td>12</td>
<td>36</td>
<td>27%</td>
</tr>
<tr>
<td>structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intensity</td>
<td>Enhanced online content compared with face-to-face content</td>
<td>15</td>
<td>4</td>
<td>6</td>
<td>25</td>
<td>19%</td>
</tr>
<tr>
<td>Student</td>
<td>Student is involved in the study because the subject-matter must be studied prior to interaction and activities</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>17</td>
<td>13%</td>
</tr>
<tr>
<td>involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulties</td>
<td>Difficulties related with schedule of the training programme, and with the newness of the subject-matter</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>11%</td>
</tr>
<tr>
<td>in the L/T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>61</td>
<td>25</td>
<td>48</td>
<td>134</td>
<td>100%</td>
</tr>
<tr>
<td>ACTIVITIES</td>
<td>Design of the activities aimed at comprehension of the theory, acquisition of useful knowledge, student’s motivation, development of attitudes, and evaluation</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>29</td>
<td>52%</td>
</tr>
<tr>
<td>Approach and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>objectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elements of</td>
<td>Inclusion of examples and clarifying elements of the activity</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>13%</td>
</tr>
<tr>
<td>structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

337
Next, some of the most important results from each of the categories are presented.

1. Theoretical content (According to students and teachers)

A total of 134 interventions or comments (100% of the interventions in this dimension) were carried out by students and teachers, in which the following sub-categories appeared: previous approach (16%), objectives (15%), elements of structure (27%), intensity (19%), student involvement (13%) and difficulties in the learning/teaching process (11%).

Of special interest are the comments from students and teachers on the structuring of the theoretical content, i.e., the resources used to reflect the theory itself. Among these elements are the objectives, the mechanisms used for the differentiation of main and secondary ideas, such as summaries, conceptual maps, etc. During DG2 the material offered was evaluated as very useful for studies, and made it easier for students to face up to the exam. This reinforces Prendes’ (2003) conception that the e-teacher should be familiar with the content in order to know how to structure, organize and present it.
“The conceptual maps helped us because to study this for the exam is unbearable ... because I imagined that when I start working in a secondary school and I have to make up a project, I’ll just come along here, look at and compare the points, and get on with the work.” (Online Student, DG1)

2. Activities (According to students and teachers)

Fifty-six interventions (100% of the total interventions of this dimension) were made, in which the sub-categories were: approach and objectives (52%), elements of structure (13%), intensity (16%), student involvement (5%), and difficulties in the teaching/learning process (14%).

The appropriate approach and definition of the objectives of the activities requires previous reflection if these are to be aimed at comprehension of the theoretical content, development of attitudes or capacities, or going even further, personal development of the students. In this sense, both teacher 2 and teacher 3 recognized throughout the interviews the ability of the activities to generate comprehension of the content both on the online and the face-to-face plane, and also to develop attitudes and modify perceptions (teacher 2) mostly on the face-to-face plane. During the interviews and again throughout DG2 they reiterated that the activities had a positive influence on students’ motivation for learning.

“What I’ve done is to eliminate part of the contents that I taught and turn it into a group activity. The objective is therefore to reflect and modify perceptions of it a little.” (Teacher 3, interview)

3. Interaction (According to students and teachers)

In this category 192 interventions (100% of the interventions in this dimension) were made, which included issues relating to approach and objectives (24%), elements of structure (27%), intensity (16%), student implication (3%) and difficulties in the teaching/learning process (14%).

Although intensity of the interactions is not outstanding in percentage, the quality of the comments led us to focus more attention on the number of contacts which took place between teachers and students. The teachers said that interaction was greater with the online students, perhaps because communication is defined by default as part of the online training system, and is part of the final evaluation (as has been reported by, among others, Stigmar & Körnefors, 2005 and Stromso et al., 2007), or because there is no pressure of timetables as there is with face-to-face classes, etc. For the teachers, this led to a relationship which was more continuous with the online students than with the face-to-face students, with the result that in some cases they came to know online students better. One teacher expressed it like this:

“Because you come here, you meet one or two people who participate more, you know their faces, etc. and that’s it: and online of course you’ve got more than one day, and if somebody intervenes again, they send you the activity, ‘I’m the one who’s in the forum’, and you relate the activity with the forum, and all that. And another will say, ‘I can’t get this activity, I can’t read it. Certainly a kind of relationship is established.” (Teacher 2, DG2)

4. Difficulties with the design (According to students and teachers)

One hundred activities (100% of those on “Design”) took place, with the following sub-categories: previous approach (17%), objectives (11%), structuring (8%), intensity (15%), student involvement (6%) and difficulties in the teaching/learning process (43%).

It is worth noting at this point that the difficulties were related fundamentally to the level of dedication that was required to follow the online course, as well as technical difficulties and those resulting from the management of the training programme. The online students described difficulties caused by the design of the programme; in DG2 they commented in particular on technical problems such as the slowness of the computers, the speed of the connections, etc. Furthermore, despite the “quality” of the training programme reported by a number of students, in some cases they felt “overwhelmed” or “saturated”, and their dedication to the course was greater than they expected, as they pointed out in DG2. The face-to-face students, for their part, spoke of difficulties which arose as the result of a course which was less complete than the online one. Even so, they admitted that they did not have to spend so much time studying. As one student pointed out in DG2:

“Basically we studied the weekend before the exam.” (Face-to-face student, DG2)
Results of the semi-structured interviews with the experts

From the interviews conducted with the experts Dr. Julio Cabero and Dr. Lorenzo García Aretio, the information necessary for analysis in this section of the investigation was obtained. A framework of categories and sub-categories was established to determine the differences between an online and a face-to-face training system. This framework is presented in Table 4, which also shows the number of comments made by the experts in each category and subcategory.

Table 4. Functions differentials between e-teachers and face-to-face teachers (experts' opinions)

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUBCATEG.</th>
<th>DESCRIPTION</th>
<th>Exp.1</th>
<th>Exp.2</th>
<th>Subtotal</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical content</td>
<td>Virtualize contents</td>
<td>Efforts to create contents for distance studying</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>Dynamize contents</td>
<td>There should not be greater dynamization in the online system</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practical content</td>
<td>Design of activities</td>
<td>Adapting the design to each system</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>21%</td>
</tr>
<tr>
<td>Interaction</td>
<td>Contact</td>
<td>In both systems contact between teacher and student is sought</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Orientation</td>
<td>There should not be greater orientation in the online system</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interaction abilities</td>
<td>Similar interaction abilities</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teacher and student involvement</td>
<td>Teachers and students seem to be more involved in the online system</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td>Psycho-pedagogical</td>
<td>The psycho-pedagogical design in the online system is better defined</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>Technical</td>
<td>The design of the techniques in the online system is important but not so much as the pedagogical design</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Totals</td>
<td></td>
<td>15</td>
<td>9</td>
<td>24</td>
<td>24</td>
<td>100%</td>
</tr>
</tbody>
</table>

1. Theoretical content (According to experts)

The experts made a total of five comments on theoretical content, 21% of the total. Of these, three addressed the virtualization of the contents in the online teaching system and two were related to the need to dynamize these contents in both teaching systems.

Both doctors stated that the virtualization of the contents of a training project requires an effort on the part of the tutor (or the content designer). The contents therefore need to be specially structured so that they can be learned autonomously. A useful option is the conceptual map for guiding the student within the conceptual context.

“When the teacher is in an e-learning context, he/she has previously made an effort to virtualize the contents, not to digitalize them, which is a different concept. Virtualization is the search for a specific structure, for which a conceptual map may exist, to present the contents in a hypertextual way until important criteria are established.” (Dr. Cabero, interview)

2. Practical content (According to the experts)

The experts made five comments, 21% of the total, on the design of the activities in the two training systems.
Just as the face-to-face teacher must choose activities that are appropriate to the learning context, it is important that the online teacher selects activities and applies strategies adapted to e-learning methodology which are not reproduced in face-to-face contexts.

“The tutor in online contexts selects activities for the students and decides how to apply strategies adapted to e-learning methodology which are not reproduced in face-to-face contexts; the application of a range of strategies is also important to avoid redundant or repetitive activities which would be tiring for the student.” (Dr. Cabero, interview)

3. Interaction processes (According to the experts)

Thirty-eight percent of the interventions addressed the interaction process in the online and face-to-face learning systems. On three occasions the experts commented on the need to establish contact between teachers and students, on guiding and coaching in two interventions, and on three occasions they spoke about the abilities necessary to interact and about how teachers and students can become involved in both training systems.

It is important to bear in mind that the aim of interaction in both systems is the same: to coach the students by taking on the role of guide and mentor. It would seem, however, that this role is being played to a greater extent in online teaching due mainly, as Dr. García Aretio pointed out, to the “distance” factor. Distance implies that the teacher has to make a special effort to foster abilities such as motivation, dynamization, and facility of learning in the students. Salmon (2004) and Volman (2005) stress the abilities of moderation and ease of learning in online activities, since they are the only way to establish a relationship with the student.

“One role that tutors and teachers must adopt is that of guide (...); evidently this is much more common in digital systems, but that doesn’t mean that in other systems there shouldn’t be more dynamism. Clearly a motivating role to facilitate learning is one that stands out today.” (Dr. García Aretio, interview)

4. Design (According to the experts)

In five interventions, 21% of the total, the experts spoke about the differences in design between the online and the face-to-face systems. On two occasions they commented on the differences in the pedagogical planning of the process, and on three occasions they referred to technical differences in design.

Technical design distinguishes a face-to-face training programme from an online one. In this respect, however, it must be noted, as Dr. García Aretio points out, that on occasions the technical design in online teaching is surmounted by the psycho-pedagogical design, resulting in failure from the perspective of training. As Ellis et al. (2006) support the design of the course can act as carrier of a good practice. According to Dr. Cabero, the systems are clearly changing in this respect, and sometimes it is difficult to establish the dividing line between the online and the face-to-face.

“The designer/manager of the process plays a bigger role in online than in face-to-face learning systems. This role is necessary in both methods, but it has become much more relevant in e-learning systems.” (Dr. García Aretio, interview).

Conclusions

In this study, the difficulties and strong points of our practices as online and face-to-face teachers have been highlighted with a view to improving teaching methods. Studies like this may induce online and face-to-face teachers to reflect on their practices, and to become aware of improvements they might make in their role as teachers.

In the first place, it has been found that the design and structure of the theoretical content of an e-training programme may, on occasions, be more satisfactory and efficacious than those of a face-to-face programme because in face-to-face programmes the teacher needs to have a previous mental structure of the contents which he/she develops in the course of the theoretical explanation. In the online mode, however, this structure is previously prepared and used as a framework for the online presentation of contents. In this sense, the use of conceptual maps and schemes, etc., to keep the student on track should be a constant feature in e-learning. Conversely, in traditional training methods, the need to guide the student through the conceptual development of the subject-matter is often neglected.
The activities in online mode may also, on occasions, be more satisfactory and efficacious than those in traditional face-to-face training. This occurs when the activities are contemplated from the design stage as the central nucleus of the learning process. In face-to-face training, conversely, the explanation of concepts often takes priority over the practice of activities. Online learners, however, may have to cope with an overwhelming amount of practical content, which may even cause them to give up their studies.

Interaction between teachers and face-to-face students can be more efficacious than with online students. Visual contact and such contact as an encouraging back-slaps, etc. are useful resources for motivating students. Online communication, too, permits enquiries to be made of the students, which can be especially valuable not only to encourage and motivate, but also to promote reflection and conceptual understanding. Nevertheless, the handling of the tools which facilitate online communication conditions the dialectic process. In both online and face-to-face systems, the consideration that interaction is part of the final evaluation conditions the student’s participation; this may have both negative consequences (students intervene without motivation) and positive consequences (students become involved in the communicative processes).

In the design of training courses, whether face-to-face or online, it is of fundamental importance that the teacher makes a previous consideration of the psycho-pedagogical theoretical grounds which will guide the teaching process. It is also important, though perhaps less crucial, that the technology and organizational support used are taken into consideration. The teachers and the institution which runs the training courses, however, tend to put more stress on online planning.

As a general conclusion to this study, no important differences are found between the functions of teachers in the two teaching modes, online and face-to-face; and if these differences do exist, they are likely to be due to the teacher’s involvement and the institution’s commitment in the programming of the learning process. In both modes, the importance of psycho-pedagogical, technical and organizational aspects of training has been shown. And the positively-valued tasks carried out by teachers are identical in both teaching systems, i.e., the facilitating of the teaching/learning process, combining the explanation of theoretical contents with activities, and encouraging interaction.

References


The Effects of Mobile Natural-science Learning Based on the 5E Learning Cycle: A Case Study

Tzu-Chien Liu¹, Hsinyi Peng², Wen-Hsuan Wu³ and Ming-Sheng Lin⁴

¹Institute of Graduate Institute of Learning & Instruction, National Central University, Taiwan // ²Institute of Education, National Chiao Tung University, Taiwan // ³Taipei Municipal Shi-Dong Elementary School, Taiwan // ⁴Department of Nature Science, Taipei Municipal University of Education, Taiwan

ABSTRACT

This study has three major purposes, including designing mobile natural-science learning activities that rest on the 5E Learning Cycle, examining the effects of these learning activities on students’ performances of learning aquatic plants, and exploring students’ perceptions toward these learning activities. A case-study method is utilized and the science club with 46 fourth-grade students is selected as the study case in the study. Besides, a set of quantitative and qualitative data were collected from the case to document the learning effects of and the students’ perceptions of the learning activities, and to discuss factors underlying these effects and students’ perceptions. The results indicate that the learning activities can enhance students’ scientific performances, including both knowledge and understanding levels. Students’ perceptions of these learning activities appear to be positive. The study identifies two factors that are prominent in the positive effects: students’ engaging in “mobile-technology supported” observation during their scientific inquiry; and students’ engaging in “mobile-technology supported” manipulation during their scientific inquiry. Finally, the conclusions that our study has drawn could constitute a useful guide for educational practitioners concerned with the potentials of mobile computing in school settings.

Keywords
5E learning cycle, Mobile learning, Natural science learning, Case study

Rationale and significance of the study

The rapid development of wireless communication and of ubiquitous-computing technology has attracted the attention of researchers from multiple disciplines. Various studies have examined the use of such new technologies in school settings in an attempt to provide learners and educators with support that is more active and more adequate (Hoppe, Joiner, Milrad, & Sharples, 2003; Hwang, Tsai, & Yang, 2008). For example, Hoppe and colleagues (2003) argue that learning with wireless mobile computing supports “active, productive, creative and collaborative learning methods” (p. 255). However, integrating innovative technologies into school settings does not necessarily ensure successful learning (Liu, 2007). As new technology emerges, some researchers are obsessed with introducing it to educational settings. These researchers endeavor to develop innovative instructional models that mirror technology functions. This approach may contribute to the innovative use of learning technologies and instructional models; however, it may also prevent effective technology diffusion owing to the neglect of teachers’ inputs. Therefore, before the large-scale implementation of mobile technology takes place in schools, educational practitioners and researchers should collaborate on effective designs based on the needs of teachers (Cobcroft, Towers, Smith, & Bruns, 2006; Liu, 2007; Peng, Su, Chou, & Tsai, 2009).

The significance of this study lies in its collaborative efforts, from the research team and the school teachers, to develop mobile natural-science learning environments and activities that are based on instructional needs as well as on an integration of mobile technology and the 5E Learning Cycle. Given its innovative features, however, the current study emphasizes the practicability of learning activities in authentic settings; therefore, instructional needs of teachers remain the priority. These results can create a balance between science-learning pedagogies (such as the 5E Learning Cycle model described in the study) and mobile learning environments and can help teachers and practitioners get educational benefits from the balance.

The 5E Learning Cycle as major scientific pedagogy

This study uses the 5E Learning Cycle as the major pedagogy herein for the following reasons. First, the 5E Learning Cycle is seen as an effective hands-on, minds-on, inquiry-based scientific pedagogy, especially for enhancing understanding (Bybee & Landes, 1988; Bybee et al., 2006; Stamp & O’Brien, 2005). Second, the 5E Learning Cycle
is one of the widely-adopted pedagogies as an indoor activity in the natural-science teaching (Bybee et al., 2006). Third, applying the 5E Learning Cycle to outdoor activities may suffer from some limitations and mobile technologies can support the application of the 5E Learning Cycle to outdoor natural-science learning.

Several studies have provided a structure that classifies various levels of inquiry-based learning (e.g., Martin-Hauser, 2002; Windschitl, 2003). For example, on the basis of the styles and levels of students’ inquiry tasks, Windschitl (2003) outlines a hierarchy of levels of inquiry, consisting of confirmation, structured inquiry, guided inquiry, and open inquiry. In general, the 5E Learning Cycle is considered as a “guided inquiry” while the teacher provides only the materials and problems to investigate and students execute their own procedures to solve the problem under the guidance of teachers (Martin-Hauser, 2002; Windschitl, 2003).

The 5E Learning Cycle, first created by Robert Karplus in the late 1950s and early 1960s, has been regarded as a general philosophy of teaching and learning with strong constructivist foundations. The 5E Learning Cycle consists of five phases: engagement, exploration, explanation, elaboration, and evaluation (Bybee & Landes, 1988; Bybee et al., 2006; Eisenkraft, 2003). It is a teaching-and-learning procedure consistent with the privileged status of inquiry and with the ways in which students naturally learn (Musheno & Lawson, 1999). The following paragraphs include explanations of both the 5E Learning Cycle phases and the major tasks for the teacher and students in each phase (Bybee & Landes, 1988; Bybee et al., 2006; Stamp & O’Brien, 2005).

**Engagement phase (E1):** The teacher assesses students’ prior knowledge and engages students in learning a new concept. The teacher also helps students make connections between prior and present knowledge, and helps to organize students’ thoughts about the learning outcomes of present activities.

**Exploration phase (E2):** The teacher provides students with a common base of activities reflective of present concepts processes, and skills. Students complete activities by using prior knowledge to generate new ideas, to explore questions and possibilities, and to execute a preliminary investigation.

**Explanation phase (E3):** The teacher focuses students’ attention on a specific aspect of their “engagement” and “exploration” experiences, and provides opportunities for students to demonstrate their understanding or skills. The teacher can also use direct instruction and guide the students toward a deeper understanding of a concept.

**Elaboration phase (E4):** The teacher challenges and extends students’ conceptual understanding and skills. Students learn to develop broader and deeper understanding and skills, through the above three phases.

**Evaluation phase (E5):** The teacher evaluates students’ progress toward achieving the instructional goals. Students learn to assess their understanding and abilities.

Several studies have shown that the 5E Learning Cycle was successfully applied to a variety of grade levels (e.g., Barman, 1992; Barman, Cohen, & Shedd, 1993; Purser & Renner, 1983; Saunders & Shepardson, 1987; Stepans, Dyche, & Beiswenger, 1988). For example, in their study of an eighth-grade genetics class, Balci, Cakiroglu, and Tekkaya (2006) compared the effectiveness of the 5E Learning Cycle with the effectiveness of expository instruction. According to the authors’ conclusions, the activities for students in the 5E Learning Cycle helped them to activate their prior knowledge and to overcome struggling with their misconceptions. In addition to the knowledge gains, these students had the opportunity to explain, to argue, and to debate their ideas, practices that helped the students further extend the conceptual understanding.

Another study applied the 5E Learning Cycle to a teacher-development program in a food-safety curriculum (Beffa-Negrini, Cohen, Laus, & McLandsborough, 2007). Seventy-one secondary teachers registered for the program, which designed food-safety training activities based on the National Science Education Standards and the Biological Science Curriculum Study. The participating teachers went through 3 modules, each with 15 hours of web-based instruction, interactive discussion, and tools for conducting experiments or for critically evaluating projects. After analyzing the 38 matched pre-test and post-test evaluations, Beffa-Negrini et al. (2007) found that the teachers were satisfied with the program and that most of them felt better able to critically evaluate food-safety information on the Internet. By experiencing the 5E Learning Cycle’s approach to learning about food safety, and by using an inquiry-based approach, these teachers could continue to effectively instruct on the same topic.
However, even given the above premises, there are certain constraints that limit teachers’ application of the 5E Learning Cycle to their natural-science teaching, especially to the outdoor inquiry-based activities. In the context of natural-science teaching, teachers often conduct outdoor inquiry, visual presentations, lab activities, or a mix of all these activities. Many educators believe that the most feasible learning activities of natural science is outdoor inquiry, which takes science investigation outside the classroom (Chen, Kao, & Sheu, 2005; Tan, Liu, & Chang, 2007). However, sometimes outdoor inquiries are more challenging and less effective than expected, owing to students’ lack of suitable resources and immediate learning supports when teachers and students are scattered in different locations (Dillon et al., 2006). Mobile computing, when combined with wireless connections, can overcome the problems and then can support students’ outdoor inquiry-based learning (e.g., Lai, Yang, Chen, Ho, & Chan, 2007; Rogers et al., 2004; Tan et al., 2007). The following section illustrates recent studies that focus on developing mobile learning environments for scientific outdoor inquiry-based learning.

Mobile computers for supporting natural-science learning

In recent years, mobile devices, wireless communications, and network technologies have emerged and have been integrated into various learning contexts (Chen, Kinshuk, Wei, & Yang, 2008; Liu et al., 2003; Soloway, Norris, Blumenfeld, Fishman, & Marx, 2001). Mobile learning environments offer much educational potential that is not easily achieved in other learning environments (Cobcroft et al., 2006; Hwang et al., 2008; Liang et al., 2005; Liu, 2007; Peng, Chou, & Chang, 2008). Mobile devices not only enable both the teacher and students to employ computing power but also do so without seriously constraining either the teacher or students regarding time or location. These constraints continue to diminish as the wireless technologies seamlessly connect mobile devices to other computing devices. This promising computing power has appealed to educators and researchers in many disciplines, particularly in the natural sciences.

Several studies have developed mobile learning environments to support scientific outdoor inquiry-based learning and have shown that these environments positively affect various learning outcomes (Chen, Kao, & Sheu, 2003; Lai et al., 2007; Rogers et al., 2004; Tan et al., 2007). For example, by using a personal digital assistant (PDA), Chen and colleagues (2003) developed a “bird-watching learning” (BWL) system to scaffold students’ learning and observational skills. In their study, the BWL system can serve as a tool for enhancing students’ curiosity, as a tool for discovery, and as a tool for independent learning. A formative evaluation of the BWL system indicated that young learners who used the system improved their learning, above and beyond what would normally be expected.

Lai and colleagues (2007) developed a “mobile-learning passport” (MLP) system to motivate and guide students’ learning in school gardens. The system can be easily operated in a wireless handheld device. The results show that students were motivated and agreed that the MLP system enhanced their learning. Rogers and colleagues (2004) focus on designing and delivering digital information to students undertaking scientific inquiry in wooded areas. Their study shows that well-organized digital information, when integrated in mobile devices, can optimize reflection and interpretation in learning. Last, Tan et al. (2007) proposed an Environment of Ubiquitous Learning with Educational Resources (EULER) based on radio-frequency identification (RFID), the Internet, ubiquitous computing, embedded systems, and database technologies to effectively present information about subjects ranging from historical relics to rare animals and geological landscapes. After the 16-week intervention, students who used EULER outperformed the control-group students regarding both post-test achievement and grades.

All of these studies showed that mobile learning environments can enhance natural-science, outdoor, inquiry-based learning. Moreover, these studies have dedicated themselves to developing various innovative learning technologies. However, most of these studies have based their developed innovative learning environment on certain learning activities and have imposed a research agenda on teachers. Under such a situation, the teachers assume little ownership in curriculum development or in teaching agenda. For example, teachers may have no opportunity to modify the mobile learning environment or learning activities to fit their own specific instructional requirements if necessary. The interest conflicts between researchers and teachers sometimes may expand, rather than close, gaps between theory and practice. The present study addresses this issue and focuses on the following related issue: teachers’ instructional/practical efforts to develop learning environments and learning activities that facilitate the effective integration of mobile technology into teaching and learning practices.
Research purposes and questions

There are three purposes of this study: (1) designing mobile natural-science learning activities that rest on the 5E Learning Cycle, (2) understanding the effects of and students’ perceptions toward these learning activities, and (3) discussing factors underlying learning effects and students’ perceptions. To achieve these goals, a mobile learning environment and learning activities are developed based on teachers’ requirements. Next, the current study collected both quantitative and qualitative data from the mobile learning activities and from related feedback of students.

The research questions are as follows:
1. What are the effects of the mobile learning activities on students’ knowledge levels and understanding levels of aquatic plants? And why?
2. What are the perceptions that the students have with regard to the learning environments and the mobile learning activities? And why?

Methods

The current study uses the case-study method. A case study is considered “an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident” (Yin, 1994, p. 13). That means, a case study is an empirical and holistic inquiry that explores a social unit, a single instance, or a phenomenon within a natural setting (Yin, 1994; Merriam, 1998). Therefore, this approach is particularly appropriate for exploring the possible effects that wireless and mobile technologies have on teaching and learning (Liu, 2007).

Participants and the context

This study took place in an elementary school in Taipei City, Taiwan, which had been constructing an ecological pool for years. With financial support from the Taipei City Government’s Department of Education, the school planned to establish a mobile learning environment devoted to natural-science instruction on aquatic plants. This plan motivated the school to collaborate with the current research team. In order to advance students’ learning of aquatic plants, the current study designed, developed, and implemented both the mobile learning activities and the learning materials in the wireless environment for a 14-week after-school science club.

The science club, consisting of a total of 46 fourth-grade students (including 24 males and 22 females), is considered the case to be studied, The students’ participation rested on their use of the ecological-pool website (described in the following section) and a tablet personal computer (tabletPC) supported by the Taipei City government. Before attending the club, these students had learned and acquired a basic understanding of aquatic plants in their science classes. Besides, they were well-informed about the instruction formats and voluntarily participated in the club. In order to examine the effects of learning activities and possible factors underlying these effects, the researchers collected and analyzed dataset regarding students’ performances, perspectives, reflections, and opinions about the instruction.

The mobile learning environment

The current study’s mobile learning environment consisted of two major components: one was an Ecological Pool website that was on the server and that connected powerful campus access points to one another; and the other constituted the tabletPCs used by each participating student. Such a learning environment enables the teacher and the students to use the resources of the website in any place (such as the ecological pool and the laboratory) at the school via the students’ tabletPCs.

In order to provide an easy delivery platform that teachers can use for their science teaching, under the support of the research team, a group of the school teachers developed the “Ecological Pool” website (in Chinese). All the information on this website echoed the instructional requests from science teachers at the school. These resources provide hands-on problem solving as well as reinforce concepts and understanding. They can be used for class
instruction, outdoor inquiry, lab activities, individual or small group study, and assessment, in a ready-made format for downloading.

Figure 1 shows the homepage of the Ecological Pool website demonstrating the actual scene of the school’s ecological pool. Users can choose a spot to enter a specific aquatic-plant section. Figure 2 presents the third section in the ecological pool. In each section, pictorials illustrate the characteristics of the plants and help students identify those characteristics. Students are able to read specific information by clicking on the portion of the pictorials or the suggested scientific terms. This function enables students to carry out inquiry without memorizing the scientific terms of the plants.

Figure 3 displays the information about the *Nuphar shimadai*: this information features illustrated collections of close-ups that correspond to each important part of a plant and that are clickable for enlarged views. The bottom-right corner provides a comprehensive reading of the *Nuphar shimadai* and shows a picture of the plant taken directly from the school’s ecological pool. The use of close-ups and of detailed information enabled students to easily locate the corresponding plant in the ecological pool. Figure 4 presents a glossary that defines the plants’ scientific terms in relation to the parts of plants, which either easily confuse students (such as a tuber and a bulb) or are not observable (such as an air-cell chamber and a lenticel). The teacher could use these materials to guide students for the compare-contrast process of identifying how these plants are alike and different. This process sometimes required students’ advanced possession of evaluating and synthesizing skills.
Figure 3. The webpage introduces the *Nuphar shimadai*

Figure 4. The glossary for the aquatic plants

Figure 5. The repository of activity sheets
Figure 5 presents a repository of learning-activity sheets, which is organized in accordance with instructional goals and with learning units. For their individual learning, students can download the activity sheets listed on the left column, while the ones on the right are designed for group work.

The mobile learning activities

In this study, seven learning activities are designed and developed based on teachers’ instructional requirements, the tenets of the 5E Learning Cycle and the school’s existing learning resources. Before the seven learning activities took place, one 40-minute technology orientation was offered to familiarize students with the mobile environment. Figure 6 shows the implementation process of these seven learning activities. As the figure shows, the five phases of the 5E Learning Cycle serve as the foundations of the seven activities. With learning devices, wireless connections, and the Ecological Pool website, the seven learning activities are executable without constraints of time or location.

In order to better illuminate the activity design, the following paragraphs describe in greater detail one activity: namely, activity six, which concerns “Submerged-type” aquatic plants. During this activity, the students learned at the ecological pool from phase E1 to half-way through phase E4, and then, they learned at the laboratory from the last half of E4 to all of E5.

Figure 6. The implementation of a mobile learning activity

Figure 7 displays specifically the implementation of activity six, which best demonstrates the use of the 5E Learning Cycle within mobile-learning contexts (refer to Appendix 1 for the procedure, the time, and the location of activity six). Figure 7 also explains the 5E Learning Cycle phases and the tasks for the teacher and the students in each phase. The instructional goals were that students should learn the specific forms and the specific habitats of the “emergent-type” aquatic plants and that the whole lesson should take 160 minutes (which included four 40-minute classes). Besides individual learning, students also have the opportunities to conduct inquiry with their team members. For instance, in the “Elaboration” phase of activity six, students work in groups to complete their activity sheets via observation, reading materials on the website, conducting experiments, and sharing findings.
Research design and data collection

In the current study, the quantitative data and the qualitative data were collected from six different data sources, as described below.

1. **A Knowledge Test:** Students took a knowledge test regarding their knowledge levels of aquatic plants (such as recall or recognition of aquatic-plant facts) before and after their science-club learning. This knowledge test used an open-ended format that required students to write down as many aquatic plants as possible and to further categorize these plants into specific types: “free-floating type,” “floating-leaf type,” “submerged-type,” and “emergent-type.” A student would receive 0 points when he or she would fail to write down the correct aquatic-plant term; a student would receive 1 point for applying a correct term to a wrong type, and 2 points for applying a correct term to its correct corresponding type.

2. **An Understanding Test:** Students took an understanding test regarding their understanding levels of aquatic plants (such as discussions, explanations, and restatements regarding aquatic plants) before and after their science-club learning. There are 24 multiple-choice items in the test (refer to the example item in the section “Results and Discussion”). The internal consistency reliability (KR-20) of the test is .71.

3. **A Learning-activity Survey** was filled out by students so that the research team could collect their overall perception of the learning activities. This survey was conducted after students finished all of the learning activities.

4. **Students’ learning activities were observed and videotaped in various instructional settings.**

5. **The students were interviewed at the end of each club meeting, and all the interviews were recorded and transcribed for later analysis.**

6. **The students kept reflective journals after each learning activity.**

After collecting data from observations, interviews, and reflective journals, the researchers coded the above datasets on the basis of attributes and dates; for instance, the reflective journal kept by student 03 on April 4th was coded as “S03-J0404,” and the observation data and interviews were coded as “O” and “I,” respectively.

Results and discussion

This section discusses results concerning (1) the effects of mobile learning activities and (2) students’ perceptions of the mobile learning activities. The presentation of these results rests on the analyses of qualitative data (such as students’ reflective journals).

The effects of mobile learning activities

In order to determine the effects of the seven mobile learning activities mentioned above, we used Bloom’s (1956) levels of cognitive learning behaviors to classify students’ learning. Students’ learning can correspond to the **Knowledge** level (recall or recognition of aquatic plant facts) or to the increasingly more complex and abstract **Understanding** level (discussions, explanations, and restatements regarding the aquatic plants).

The knowledge test

The present study performed a *t*-test to compare the mean scores that students received in the knowledge test before and after students’ engagement in learning activities. The results in Table 1 show that students’ mean scores (26.04) after learning activities were significantly higher than students’ mean scores (10.30) before learning activities (*p* < .001). This result confirms that students’ knowledge of aquatic plants increased and that the students could better identify correct aquatic-plant terms and the plants’ corresponding types after the mobile learning activities.

| **Table 1. Students’ knowledge test scores regarding aquatic plants before and after learning activities (N=46)** |
|--------------------------------------------------|----------------|---------------|----------------|---------------|----------------|
| **Before learning activities** | **After learning activities** | **T** |
| **Knowledge test of aquatic plants** | Mean | SD | Mean | SD | *** |
| **Mean** | 10.30 | .87 | 26.04 | 1.81 | 11.483 *** |

*** *p* < .001
The understanding test

Results of the t-test in Table 2 indicate that students’ scores after learning activities (M=19.04) were significantly higher than students’ test scores before learning activities (M=15.63) (p < .001) in the understanding test on aquatic plants. This analysis result means that students’ understanding of aquatic plants increased after their engagement in the learning activities.

**Table 2. Students’ understanding test scores regarding aquatic plants before and after learning activities (N=46)**

<table>
<thead>
<tr>
<th></th>
<th>Before learning activities</th>
<th>After learning activities</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Understanding test on aquatic plants</td>
<td>15.63</td>
<td>3.41</td>
<td>19.04</td>
</tr>
</tbody>
</table>

*** p < .001

Possible factors underlying the effects of mobile learning activities

According to the 5E Learning Cycle, students’ use of a hands-on inquiry process enhances not only the students’ understanding but also shifts in the students’ understanding (conceptual changes) (Bybee & Landes, 1988; Bybee et al., 2006). In the past, this pedagogy was often adopted as indoor learning activities (such as in a lab) (e.g., Barman, 1992; Barman, Cohen, & Shedd, 1993; Purser & Renner, 1983; Saunders & Shepardson, 1987; Stepans et al., 1988); yet the learning activities in this study have combined mobile learning environments developed by school teachers and the tenets of the 5E Learning Cycle, the goal being to help students conduct scientific-inquiry activities in various learning settings. Of these learning activities, mobile computing and wireless connection, owing to their convenience and immediacy, have made possible a flexible transition among various learning environments (such as an ecological pool, a science laboratory, and the Ecological Pool website). Each learning environment, in turn, would have its own distinctive resources to which the students could flexibly and easily adapt during the mobile learning activities. For instance, students could gain both micro and macro views of a water lily by combining a lab setting and an observation activity at the ecological pool. In the lab setting, students examined the air-cell chamber of water lilies through electronic microscopes; while at the ecological pool, the teacher guided the students as they explored the habitats, the relationships between the air-cell chamber and habitats, and the roles that water lilies played in the whole ecological system, from a macro perspective in the mobile environment. The following section treats the factors that might have resulted in positive learning effects, and does so on the basis of qualitative data collected from observation, students’ learning journals, and interviews.

First, students’ efforts to observe aquatic plants by using adequate support of mobile devices might have helped the students correct their related misconceptions, in turn, strengthening the students’ understanding of aquatic matters. While learning in these mobile learning activities, students had opportunities to observe or to manipulate scientific resources. The data collected from students’ reflective journals and from interviews confirm that the use of resources and the guidance on the Ecological Pool website had positive effects on student learning. For example, students (11, 15, and 22) were able to identify the stems and stalks of “free-floating type” aquatic plants after reading articles on the website and observing the plants. Another example took place in the activity “Water lily, lotus, pond lily.” Before this activity, most students were confused about the characteristics that distinguish lotuses from water lilies, and some students even believed that the two were the same type of plant. In this activity, students carried out their inquiry guided by the website, and they could successfully identify the differences between lotuses and water lilies afterwards. Student 37 and her teammate discovered leaf differences (hairy thorn-like growths on lotuses, and sawtooth-like leaves on water lilies) (J0609), and student 45 and her teammate concluded that the lotus is categorized as an “emergent type” of aquatic plant, while the water lily is categorized as a “floating-leaf type” aquatic plant (J0609).

One factor underlying positive effects of knowledge acquisition and conceptual change lies in students’ hands-on experiences in the learning activities. For example, in the fourth activity of “free-floating type” aquatic plants, students used electronic microscopes and projected images of the cilia and the air-cell chambers of certain aquatic plants onto the electronic whiteboard. This experience provided the students a common base of activities to identify current concepts, misconceptions, processes, and skills. Writing in the reflective journal, student 34 identified one of his misconceptions: “This activity challenged my previous assumption about salvinia natans. The hundreds of air-cell chambers on the surface of salvinia natans really shocked me because, before the activity, I supposed there were
Another factor underlying positive learning effects might be students’ prolonged engagement in the mobile learning activities. The teacher discovered that students loved the observation activities at the ecological pool and they loved to share their findings after the inquiry. Moreover, instances of these trends received frequent mention in students’ reflective journals: for example, student 21 wrote, “These tasks kept me very busy, but I really enjoyed this unique learning experience!” (J0510). Student 22 mentioned, “Even though there were many mosquito bites on my legs, I still felt a sense of accomplishment after figuring out the answers through the use of the website and through the observation at the pool” (J0510), while student 09 characterized his inquiry as a “joyful experience of treasure hunting” (J0607). More detailed discussions of students’ feedback are presented in the section “Students’ perceptions of the mobile learning activity.”

Having discussed the above factors’ role in positive learning effects, this study acknowledges limitations in its activity design and in its implementation, and does so on the basis of students’ understanding tests, interviews, and observation logs.

First, even though these mobile learning activities heavily emphasized observation during students’ inquiry process, specimens of several plant categories, such as lotus roots, lotus seeds, and lotus seedpods, were not available for observation during the research period. Therefore, students could learn about these categories only through readings posted on the website. The lack of in-depth observations resulted in a low correct rate (59%), which was significantly lower than the average correct rate (79%), in one of the understanding-test items: “What is the function of the ‘holes’ in the lotus root?” In-depth interviews also showed that students retained their misconceptions after the learning activities. For example, two students (student 17 and student 34) who had chosen the wrong answer of option 2 (As holders for lotus seeds) conjectured that the round shape of those holes would match that of lotus seeds. These students’ responses might have derived from their inability to observe the growth of these plants. These findings suggest the importance of matching the curriculum to seasonal patterns, in order to promote students’ learning by observation.

Next, although the mobile learning environment, developed by the team of school teachers, can meet teachers’ instructional needs, and it is easy to operate and maintain, it might be with somewhat limited technical stability. For instance, in the early stages (activity 1) of this study’s activities, certain technical problems (e.g., disconnection) hampered the processes of learning. Although the solutions of the technical problems were soon identified—solutions including the promotion of student tutors and better preparation before the activities, the technical problems still may negatively influence students’ learning.

Students’ perceptions of the mobile learning activities

After collecting the data from the learning-activity surveys and the after-project interviews regarding students’ experiences in the learning activities, we described in detail (1) the students’ overall impressions of learning activities and (2) the students’ perceptions of the mobile devices and (3) the students’ perceptions of the learning activities. The following subsections present these detailed descriptions.

Students’ overall impressions

The results indicate that students’ overall impressions of the activities were positive: twenty-nine (64%) out of forty-five students who completed a Learning-activity Survey expressed that they greatly liked or liked the activities, and sixteen (36%) students expressed that they moderately liked the activities. We then analyzed the students’ motives that had been associated with the students’ aptitudes.

The current study investigated students’ motives by providing the students with a survey question for which they could select one or more predefined options (see Table 3). Twenty-two students (49%) felt that their opportunity to use a tabletPC throughout the learning activities was a source of motivation. Sixteen students (36%) responded that their being able to observe the “real” aquatic plants in the ecological pool was essential, while one-third of the
students felt that they had gained knowledge of aquatic plants and that the activity design and the teacher’s delivery of the activities were interesting.

<table>
<thead>
<tr>
<th>Motive</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was motivated by my opportunity to use a tabletPC.</td>
<td>22</td>
</tr>
<tr>
<td>I could observe the aquatic plants in the ecology garden.</td>
<td>16</td>
</tr>
<tr>
<td>I think I have gained knowledge of aquatic plants in the learning activities.</td>
<td>15</td>
</tr>
<tr>
<td>I felt that the learning activities were interesting.</td>
<td>14</td>
</tr>
<tr>
<td>I enjoyed the way my teacher taught in this class.</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3. Students’ overall impressions of the learning activities (N=45)

Students' perceptions of the mobile devices

Because the tabletPCs had played a vital role in the class activities, we investigated students’ perceptions of the mobile-learning devices by collecting and analyzing students’ reflective journals and in-depth interviews. Some students had viewed their tabletPCs as a tool (e.g., a reference tool) for studying aquatic plants (students 07, 08, 09, and 24), and other students even had regarded it as a handy dictionary for all plants (students 17 and 41). For example, student 09 stated, “The tabletPC is very important to our class activity. I can look up a specific aquatic plant via the website and display the features when I’m at the ecological pool; also, the information that comes up can help me understand why a specific plant is unique” (I0626). Another student (student 41) mentioned that the features of a tabletPC are akin to those of a dictionary: “If we had a hard time locating a certain aquatic plant, we could always search for online information using the tabletPC” (I0626).

Students' preferences regarding the mobile learning activities and lectures

From students’ reflective journals, the researchers were able to understand students’ perceptions and preferences regarding the two instructional modes (the mobile learning activities in the current study and the lectures that students often take in ordinary classrooms). The results show that, of the forty-three students who completed the journals, twenty-four students (56%) preferred the mobile learning activities to lectures, nine students (21%) liked both of the modes, and ten students (23%) preferred lectures. The following sections describe in detail students’ preferences and feedback.

Why students preferred mobile learning activities. Some students variously characterized the mobile learning activities as more “personal,” “collaborative,” and “hands-on” than were the lectures. For example, students 10, 30, and 35 found the mobile learning activities more “interesting and interactive” than the teacher’s lectures. Student 10 stated, “The mobile learning activities help me appreciate the hands-on experiences through observation” (J0524), and students 30 and 35 raised similar points. For example, student 30 stated, “We prefer learning with our teammates and having an inquiry-type learning experience. We’ve built a close relationship by solving problems and completing our assignments together” (J0526). Some students believed that they had gained knowledge not by memorizing facts about plants, but by observing authentic aquatic plants (students 07, 09, and 39), and one student, student 28, mentioned that he sometimes felt like a researcher conducting scientific inquiry and that this feeling motivated him to search for resources beyond class requirements (J0526).

Why students preferred both mobile learning activities and lectures. Some students liked both mobile learning activities and lectures. For example, student 29 said, “The lectures can help me understand the concepts; however, the mobile learning activities can help build up thinking skills” (J0524). Student 06 believed that “I can gain knowledge quickly from lectures, while I can feel a sense of accomplishment by looking for resources to solve problems myself” (J0524). Student 11 pointed out, “Lectures are good sources of knowledge, as mobile learning activities create a relaxed learning atmosphere for collaboration.” As such, some students suggested an ultimate learning arrangement where a combination of these two modes can provide an adequate mixture of independent and collaborative learning experiences for students with diverse learning styles (e.g., S06-J0524, S17-J0524, and S40-J0526).
Why students preferred lectures. Students 13 and 20 mentioned similar points, with student 13 declaring, “The inquiry-type learning is very time consuming. Sometimes my teammates come up with different solutions, and we’re not even sure whether or not these are the correct ones. This made me feel as though I’d spent my time inefficiently; therefore, I prefer the direct instruction from the teacher—it’s much easier and right-to-the-point” (J0524). Student 01 and 03 echoed the above statement. Student 01 said, “It’s easier to catch the key points from the teacher’s lecture … so that I can finish the activity sheet promptly.” Student 03 pointed out, “I liked lectures because they’re easier” (J0524).

The results revealed that most students’ perceptions of mobile learning activities were very positive because the students viewed their mobile-learning devices as an integral part of their learning. It was apparent that most students preferred mobile learning activities to lectures because of such activities’ ability to promote hands-on tasks, observation, and inquiry-based experiences, whose characteristics have matched the important characteristics of mobile learning activities. Even so, it is worth noting that there were still a few students not ready to the learning strategies (e.g., hands-on inquiry or observation) required by the mobile learning activities. For that reason, in order to accommodate students’ learning styles, the teacher can deliver essential knowledge via lectures and can introduce the more hands-on and exploratory modes of mobile learning activities.

Conclusions

This study designed and developed mobile natural-science learning activities based on teachers’ instructional requirements, the tenets of the 5E Learning Cycle, and the school’s existing learning resources. The purposes of this study were to report the learning effects of and students’ perceptions toward learning activities that rest on the tenets of the 5E Learning Cycle and possible factors underlying these positive effects and students’ perceptions. From our findings, we have drawn the following conclusions and suggestions for educational practitioners and researchers regarding the integration of mobile computing into future teaching and learning settings.

First, the present study’s learning environment consisted of one computer, several wireless access points, a website containing learning materials and activity sheets, and several mobile devices for students. The study has shown us that constraints on time and on location diminish as technology seamlessly interweaves scientific-learning pedagogies (such as the 5E Learning Cycle model) with mobile learning environments, especially for students’ observation activities at the ecological pools. Such findings echo assertions that the mobile computing, combined with wireless connections, can overcome problems arising during students’ outdoor activities and can enhance the quality of outdoor, natural-science, inquiry-based learning (Lai et al., 2007; Rogers et al., 2004). As more innovative features mobile technology become available, however, not all of them fit instructional goals or learning needs in the school settings. For example, students use the global positioning system (GPS) to locate and gather information at various locations in some studies. Yet in the current study, until this point, there seems no need to involve GPS because of its limited size of the ecological pool as well as its difficulties in locating moving targets (e.g., plants) on the pool. The complexity of developing and maintaining such a system might have affected teachers’ aversion to use mobile technology. Also, we believe that, via the compare-contrast process of identifying plants guided by the website, students can gain deeper levels of knowledge and understanding of the aquatic plants than merely following the directions of GPS. Therefore, when introducing technology into classrooms, researchers should consider the potential of the technology that, for most educators and practitioners, is more useful, easier to use and more accessible in their daily practices.

Second, we have shown that the 5E Learning Cycle model is an effective hands-on and inquiry-based scientific pedagogy that enhances students’ scientific knowledge and understanding. When combined with mobile computing, the model seems to create opportunities in which teachers can design unique learning experiences that transition smoothly from indoor inquiries to outdoor inquiries. The results of the study suggest that, in general, the mobile learning activities based the 5E Learning Cycle model can effectively increase students’ knowledge and understanding of aquatic plants and improve students’ learning motivation. Future research and teaching would do well to reference this best-practice scenario, in which we detail the design and the procedures of the 5E Learning Cycle phases as well as the roles of, and the tasks for, the teacher and the students in each phase in the mobile learning contexts.
Third, the current study’s case presents an in-depth view of learning in the mobile learning activities that rested on the 5E Learning Cycle; however, the study’s use of a case may also limit the generalizability of the study’s findings. Based on the analysis of learning effects and factors already discussed here, future studies may extend the current study’s findings to other classes or disciplines, or may use alternative research methods (such as field-experiment methods) to present a more thorough picture of effective integration of mobile technology into teaching and learning practices.

Acknowledgement
This study was supported by the National Science Council (NSC) of the Republic of China (Taiwan), under grant numbers NSC 96-2520-S-008 -007 -MY2 and NSC 97-2631-S-008-003. Thanks are given to: Taipei Municipal Dahu Elementary School, and Taipei City Government Educational Department Informational Education.

References


### Appendix 1. The phases and the procedures of the 5E Learning Cycle for Activity Six

<table>
<thead>
<tr>
<th>Phase</th>
<th>Procedure</th>
<th>Time (Minutes)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Engagement (E1)</strong></td>
<td>The teacher assesses students’ prior knowledge and engages students in a new concept by having them browse the Ecological Pool website. Students familiarize themselves with the thirteen emergent-type aquatic plants and fill out the pre-activity sheet online. The teacher also helps them connect their past and present knowledge.</td>
<td>15</td>
<td>Ecological pool</td>
</tr>
<tr>
<td><strong>Exploration (E2)</strong></td>
<td>Students carry out their preliminary investigation of “emergent-type” aquatic plants by filling out the activity sheet via observation at the ecological pool or search for information on the website. The teacher provides students with a common base of activities within concepts discussed at the Engagement phase.</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td><strong>Explanation (E3)</strong></td>
<td>The teacher guides the discussion on the activity sheet and helps students locate the exact spots of these thirteen emergent-type aquatic plants.</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td><strong>Elaboration (E4)</strong></td>
<td>Working in groups of two, students refine their activity sheets through observation and access the website. Students are encouraged to demonstrate and explain their understanding of the “emergent-type” aquatic plants. Working in groups to incise the leafstalk of plants. They fill out the activity sheet online after conducting research on the question “Do roots of aquatic plants decay in water?” Students watch the video “The World of Water Plants” to reinforce their learning.</td>
<td>40</td>
<td>Science laboratory</td>
</tr>
<tr>
<td><strong>Evaluation (E5)</strong></td>
<td>The teacher evaluates students’ activity sheets. Students learn to assess their understanding via group discussion.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Integration and Deployment of Educational Games in e-Learning Environments: The Learning Object Model Meets Educational Gaming

Javier Torrente, Pablo Moreno-Ger, Iván Martínez-Ortiz and Baltasar Fernandez-Manjon

Dpto. Ingeniería del Software e Inteligencia Artificial, Facultad de Informática, Universidad Complutense de Madrid, Spain // jtorrente@fdi.ucm.es // pablom@fdi.ucm.es // imartinez@fdi.ucm.es // balta@fdi.ucm.es

ABSTRACT
Game-based learning is becoming popular in the academic discussion of Learning Technologies. However, even though the educational potential of games has been thoroughly discussed in the literature, the integration of the games into educational processes and how to efficiently deliver the games to the students are still open questions. This paper addresses the aspects of integration and automatic deployment of educational games in Learning Management Systems. This integration simplifies the introduction of games in educational settings, leveraging the pre-existing technological infrastructure. Our approach is based on the automatic packaging and exportation of games as self-contained Learning Objects that can be easily distributed through any LMS compliant with the current interoperability standards. We thus inherit the advantages of the Learning Object model in terms of interoperability and, when supported by the LMS, in terms of student tracking and assessment.

Keywords
Educational games, Learning Objects model, Learning Objects meta-data, Learning Management System interoperability

Introduction
Ever since videogames were born, they have been identified as potential learning tools (Malone 1981b; Malone 1981a). Since then, there has been a sustained academic discussion of their educational applicability (Leutner 1993; Rieber 1996; Papert 1998; Prensky 2001; Gee 2003; Mitchell and Savill-Smith 2004; Squire 2005; Michael and Chen 2006). This discussion has grown and evolved along with the maturity of the videogame industry, which has now become one of the most relevant businesses in the entertainment sector. After years of discussion and debate, the benefits of applying digital games, although lacking enough proof of the results sometimes (Hays 2005), have at least caught the attention of a great part of the academic community (Van Eck 2006). However, this does not mean that we are ready to massively adopt game-based learning in the educational system. The discussion has now moved on to other aspects such as pedagogically sound game design, student tracking and assessment, cost-effectiveness, and integration aspects (de Freitas and Oliver 2006; Van Eck 2006; Burgos et al. 2007b; Torrente et al. 2008a).

In this work, we will not be focusing on the strictly pedagogical aspects of game-based learning (referring interested readers to detailed literature reviews on the topic), but on some technical issues perceived as a barrier in the adoption of educational games that are hindering further research in this area. Two of these issues are how the games are delivered to the students and how to track the performance of the students in the game experience. A possible approach to tackle these issues is the integration of the games as a new type of content in e-Learning systems, thus taking advantage of the current widespread infrastructure provided by the modern Learning Management Systems (LMS). However, commercial videogames are difficult to deliver through these systems mostly due to their significant technological requirements (e.g. specific operative system or specific hardware) and due to an inherently complex distribution process. On the one hand, instructors need to address the packaging and distribution of games and, on the other hand, the students need to obtain the game, install it and then execute the game on their computers, which need to meet the technical requirements.

In this regard, the <e-Adventure> platform overcomes the technical issues related to installation, distribution and platform-dependency by providing an educational game engine that can be distributed via web (Moreno-Ger et al. 2008b) through the use of Java™ technologies. Although, the Java-based approach alleviates the interoperability and execution issues, it does not adequately address all the distribution issues that instructors have to face. The instructor still needs to cope with the integration in the educational process. For this reason we propose an approach that takes advantage of pre-existing e-Learning environments. This approach relies on the production of educational
videogames as self-contained distributable Learning Objects (LO) following the principles of the LO model (Polsani 2003; Balatsoukas et al. 2008). This model envisions the learning contents as small self-contained objects that can then be combined in larger units (i.e. courses). In addition, there are standardized formats to store and distribute learning objects, allowing the content to be reused across platforms and contexts. This standardization guarantees the simplicity of the deployment of these contents in heterogeneous LMS.

Therefore the paper is structured as follows: firstly we discuss aspects related to the domains of e-Learning and educational gaming, as well as present some of the issues concerning their integration. Then we contextualize this contribution by analyzing related work. Finally, we describe the <e-Adventure> platform and how it facilitates the aforementioned integration, illustrating the process with a case study in which we test and analyze the integration of an <e-Adventure> game in some commercial and open source LMS.

E-Learning and Videogames

In this section we provide a general overview of both e-learning and educational gaming and discuss some issues that need to be addressed to simplify the use of games in education.

Current trends in online learning: developing standards and specifications to promote content interoperability

The e-Learning field is an industry on the rise that is beginning to be considered a mature technology. The initial excitement about the possibility of accessing content anytime and anywhere was hindered by simplistic e-learning systems that were essentially repositories with vast amounts of content and very basic facilities (e.g. management or communication) (Weaver 2002; Zemsky and Massy 2004). However, modern e-Learning systems are more comprehensive and try to mitigate the problems of the separation between students and instructors, such as the lack of motivation or the high drop-out rates. Modern Virtual Learning Environments, often labelled as Learning Management Systems (LMS), provide facilities for the interaction between instructors and students, detailed tracking of the students’ progress, and a simple path for the delivery of content through the web. In addition, their use is getting more and more generalized in diverse contexts, not only as an alternative to face-to-face learning, but also as a rich complement (e.g. most universities are using an LMS to complement and support their lectures). This approach has been called b-Learning (Osguthorpe and Graham 2003; Garrison and Kanuka 2004).

However, as e-Learning has become an important aspect of many learning experiences, there is a broad range of competing platforms. Additionally, we cannot simply throw any kind of content at the students and expect them to learn. There is a need for high quality content, built with solid educational principles. This means that the authoring and maintenance costs for this content are becoming huge, and the variety of competing platforms may put the investment at risk if those expensive contents are not interoperable.

The Learning Object (LO) model (Polsani 2003; Balatsoukas et al. 2008) addresses these issues by proposing a development strategy of learning content based on self-contained pieces that can be assembled into courses, supported by standardized interchange formats to simplify the interoperability of contents among systems and avoid vendor lock-in. With this objective, the e-Learning arena has been immersed in a process of standardization and specifications development to support the interoperability between diverse LMS and content repositories.

Thus far, some specifications and standards are gaining acceptance in the e-learning market. Regarding the encapsulation of content, the effort carried out by the IMS Global Consortium has achieved a high impact with the IMS Content Packaging (IMS CP) specification (IMS Global Consortium 2004). The specification establishes a standardized format for the packaging and distribution of LO. According to the IMS CP specification, contents should be packaged in a zip file containing all the learning contents, along with a manifest file which provides information about the structure of the learning contents and additional information about how to deploy and deliver the content. Most of the commonly used LMS have facilities to import and export IMS CP contents, such as Moodle (Dougiamas and Taylor 2003), Sakai (Farmer and Dolphin 2005) or WebCT/BlackBoard (Goldberg and Salari 1997). This widespread adoption suggests that IMS CP can be taken as a preferred standard when it comes to packaging content.
In addition, the IMS CP specification is flexible and can be customized to specific scenarios through the so-called Application Profiles. One such profile is the Shareable Content Object Reference Model (SCORM) Content Aggregation Model (ADL 2006), created in the context of the Advanced Distributed Learning (ADL) initiative.

ADL SCORM not only covers the packaging of learning objects, but also provides a communication protocol between an LMS and the Learning Objects. This communication protocol allows the LMS to gather tracking and assessment information generated within the LO. In addition, the latest version of the ADL SCORM reference model introduces the concepts of Sequencing and Navigation (SCORM SN). SCORM SN allows content developers to create activity sequences and to define the interaction mechanisms to navigate through them. Thereby the interaction between the student and an LO can affect the sequencing process through the aforementioned communication mechanism (Gonzalez-Barbone and Anido-Rifon 2008).

This notion of having complex learning sequences that vary depending on the outcomes of the individual activities is also present in another IMS specification: IMS Learning Design (IMS LD) (IMS Global Consortium 2003). In IMS LD, the LOs are part of the environments provided to the student during the exposition of activities and their outcomes can affect future branching decisions during the learning experience.

On the other hand, in order to address the aspects of storage, search and retrieval of LOs in content repositories, content developers are encouraged to describe and annotate with metadata their LOs (Anido-Rifón et al. 2002). This annotation process also requires a standardization effort in order to define how the objects should be described. One of the most commonly used metadata specifications in this context is the IEEE Learning Object Meta-Data standard (IEEE 2002). This standard specifies a wide range of metadata that can be used to describe LO characteristics. The data is organized in categories concerning diverse aspects (general data, technical and educational features, etc.).

The potential of videogames for education

The academic interest in studying the educational potential of computer and videogames has rapidly increased in the last few years. The trend started with some seminal works about the relation between the motivation and engagement provided by games and learning, such as (Malone 1981a) or (Lepper and Cordova 1992). From these works, many authors have highlighted how games can keep people of different conditions focused and concentrated on a task during long periods of time. In fact, further studies about what make learning fun have identified aspects that are aligned with learning principles (Jenkins et al. 2003; Aldrich 2004; Becker 2007; Gee 2007), favoring the development of skills and competences rather than fact memorization.

However, as some recent meta-analyses and literature reviews (Hays 2005; Dondlinger 2007) reveal, the conclusions of the research on game-based learning are disparate. Although multiple authors identify signs to argue in favor of the benefits of this approach, its effectiveness has not been demonstrated thoroughly yet. Some researchers reached no significant conclusions, others could not prove game-based learning to be better than other instructional approaches, and others are sometimes questioned because of the methods used in their experiments (Hays 2005).

However, even if this line of research is still under discussion, there are enough evidences and cases of successful (past and present) applications of instructional videogames and game-like simulations in real contexts, which indicates a high interest in this approach (Aldrich 2004; Lunce 2006; Michael and Chen 2006; de Freitas and Jarvis 2007). From this perspective, the main objective of this work is not to further elaborate on the pedagogical discussion, but to study some of the technical issues that represent a barrier for the introduction of game-based learning initiatives in our educational processes. Reducing these barriers should facilitate further research about the specific educational values of game-based learning.

The problem of delivering the games

One of the main barriers that hinder the adoption of educational games is the complexity that they introduce in the learning process. While a lecture does not require any technology investment, videogames require up-to-date computers and controlled environments. Schools usually lack this kind of equipment: In a literature review on game-based learning presented in 2004, Kirriemur and McFarlane (2004) listed a number of concerns impeding the
development of educational games including, among others, “the almost total lack of video gaming equipment in schools, as opposed to homes”. Moreover, even when this kind of equipment is present in the schools, they usually lack the staff preparation and/or the time required to organize educational gaming sessions. Similarly, Rosas et al. (2003) identified as main issues preventing the use of game-based learning the resistance of teachers towards new technologies and the complexities of edutainment along with, again, the lack of technological infrastructure.

To tackle this aspect, an effective mechanism to distribute the games to the students would be to deliver them through a LMS. Delivering a game through a LMS is more effective than other approaches, like handing out CD-ROMs to the students in class or arranging in-school game sessions. Additionally, the tracking and assessment features of some modern LMS would give the instructor more power to control aspects of the learning experience such as when the students accessed the game, the time each student spent playing it and the outcomes of the game session. However, deploying a game in a LMS is still a complex task that demands specialized knowledge from instructors, as previously mentioned.

Another key issue is the philosophical difference between videogames and modern LMS, two completely different industries. On the one hand, e-Learning technologies have been developed around the web following a philosophy of interoperability between platforms, simplicity for the instructors, making contents available anytime and anywhere, etc. Modern LMS have been devised for the distribution of online courses that are mainly based on web contents (not only html documents but also richer formats such as PDF or multimedia resources) that are packaged and distributed according to a set of standards and specifications, as it was mentioned in the previous section.

On the other hand, commercial videogames are usually one of the most resource-consuming products of the software industry. As a proof of this claim, the videogame industry is pointed out because of the key role it plays in the promotion of technology (both hardware and software) innovation (Crandall and Sidak 2006). This fact, along with the inherently high development cost of commercial games (currently averaging in the range of $10M-$25M), forces developers to follow a design strategy in which runtime performance is a key aspect. Hence videogames are usually closed products, difficult to deploy in heterogeneous environments (videogames are rarely cross-platform), and hardly ever designed to be deployed directly from the web. That makes their integration in current LMS a big challenge.

Considering this, it is mandatory to find solutions for the development of educational videogames that can be easily integrated in online systems, and that follow the current trends in standardization described previously. For that purpose, we propose a simple methodology for the development of educational videogames targeted to web-based learning environments composed of two steps: First, we need to tackle the technological barrier between the e-Learning and videogames fields (i.e. produce web-oriented videogames). Then, we need to package, describe and deliver the games using the same standards and specifications that are currently being used in those web-based learning environments.

Related work

Game-based learning is a very broad field, with varied initiatives and heterogeneous approaches. An exhaustive analysis is beyond the scope of this work, but we will compare our approach with some initiatives that we consider especially relevant to the e-learning domain because they specifically address authoring or deployment issues.

Tools for the creation of educational videogames

Focusing on the instructors who demand the ability to create and modify their own games to suit their specific needs (Torrente et al. 2008b), there are some author-friendly environments specially oriented to facilitate the construction of videogames. Some examples are the initiatives led by The Game Creators (http://www.thegamecreators.com) a company that provides commercial authoring tools for the creation of all kinds of videogames at a reasonable cost. Perhaps their two most representative products are The FPS Creator (http://www.fpscreator.com) (an environment for the creation of First-Person-Shooter games with amazing results) and The 3D Game Maker (http://t3dgm.thegamecreators.com). In this line, most products are focused on a single game genre which simplifies both the cost of the tool and the complexity of its use. Probably the most represented genres are adventure games and
interactive fiction, represented by tools like Adventure Maker (http://www.adventuremaker.com) or ADRIFT (http://www.adrift.org.uk/cgi/new/adrift.cgi). Some examples of more complex tools (probably beyond the skill level of most instructors, but which still simpler than fully-featured programming environments) would be projects like Mission Maker (http://www.immersiveeducation.com/MissionMaker) or Alice (http://www.alice.org), both specifically purposed for education.

**Integrating games into learning experiences**

There is also some existing research about how to deliver educational videogames and game-based simulations. Given that most studies are focused on studying the pedagogical aspects, few works actually reflect on the mechanisms for delivering the games to the students.

The problem of delivering the games is, in fact, closely related to the integration of these games with other learning contents (traditional materials or other educational games) in bigger courses. In this line, some interesting approaches have been proposed, relying on game-based environments (which are usually friendly for students) not only for playing, but also as interfaces for browsing other materials (Chao 2001; Christoffel and Schmitt 2002). For instance, Quest Atlantis (Barab et al. 2005) is a 3D multi-user environment mostly used to immerse children (ages 9-15) combining educational gaming with other lessons defined by teachers. The environment supports large courses in the form of quests that the student must accomplish by carrying out diverse tasks (game-based and not game-based) without leaving the game world. River City (Ketelhut et al. 2006) follows a similar pattern. However, these approaches are self-contained and are not focused on the reuse of the learning content produced.

Focusing on the delivery of games as Learning Objects, the proposal in (Burgos et al. 2007b) contemplates the inclusion of the games in IMS Learning Design-based Units of Learning (Koper and Tattersall 2005), which is a very similar approach to our proposal. However, it does not contemplate the encapsulation of the games themselves as self-contained Learning Objects for their direct deployment in e-Learning environments or their individual storage in content repositories. The games are described as activities within the Unit of Learning. Thus these works should be complemented with the proposal of more fine-grained packaging and annotation mechanisms for individual games, which should facilitate the actual integration of the games into those game-based Units of Learning. In addition both works are also complementary in terms of enabling the games to affect the behaviour of the learning experience. In their work, Burgos et al. identify as a key unsupported feature the possibility of allowing the outcomes of the games to affect the inner state of the Unit of Learning that is used to make branching decisions.

**Production of educational videogames as Learning Objects: the <e-Adventure> platform**

The <e-Adventure> platform is an environment for the development of educational point-and-click adventure videogames that addresses the technical and distribution issues described previously, covering both the development (and maintenance) of the games and their execution (Moreno-Ger et al. 2008a). As stated previously, we propose to develop educational videogames as web content in order to facilitate the delivery of the games to the students through online LMS. To facilitate this process, the <e-Adventure> platform was developed using Java technologies, which were designed bearing in mind their application in web environments.

In order to reduce the technical requirements of the games and retain both student interest and educational value, the genre of point-and-click adventure games was chosen. The low technical complexity of these games does not imply an excessive reduction in the advantages of using videogames in education, as the genre supports all the features that are relevant from an educational point of view and still manages to provide a fun approach to learning (Ju and Wagner 1997; Van Eck 2007). Besides, narration is a key aspect of adventure games, which can be used effectively from an instructional point of view (Dickey 2006).

**Development of <e-Adventure> games as Learning Objects**

However, as previously mentioned, the creation of web-deliverable learning objects is just one half of the problem. Game-based content is complex and expensive to produce, and this increases the need of interoperability between different learning platforms. For this reason, <e-Adventure> games can be exported as self-contained Learning Objects (LO) following the IMS Content Packaging specification (IMS Global Consortium 2004). The process is
straightforward: instructors use the <e-Adventure> editor to create and modify the game and then the editor offers the possibility of exporting the executable game as a LO.

Additionally, following the notion of active LOs that exchange information with the LMS for tracking and assessment purposes, <e-Adventure> games can communicate at runtime with the LMS host. Whenever an <e-Adventure> game is launched, the game engine tests if it is possible to communicate with the host LMS (Martinez-Ortiz et al. 2006). If the test is successful, the game engine periodically communicates with the LMS notifying any value update of a set of educational properties defined during the creation of the game by the instructor. If no communication link can be established, the engine works in a disconnected mode.

<e-Adventure> games support two protocols for the communication with the LMS host. The first approach is to use the IEEE ECMAScript API for Content to Runtime Services Communication (Richards 2004). If this protocol is not supported by the LMS, the engine also tries to establish a link using an ad hoc protocol developed as part of the <e-Adventure> platform. The standardized approach uses the same set of specifications as the ADL SCORM reference model, thus being compatible with SCORM-compliant LMS. On the other hand, the proprietary protocol was developed to enable the communication with IMS Learning Design (IMS Global Consortium 2003) environments, and has been successfully tested with the CopperCore IMS LD (Burgos et al. 2007a) and the .LRN platform (Moreno-Ger et al. 2008a).

The aspect of allowing the annotation of the LO with technical and educational metadata is covered according to the IEEE Learning Object Metadata (IEEE 2002) standard. Due to the scope of the <e-Adventure> editor, the objective was not to provide a full IEEE LOM editor. Instead, we have focused on a minimal subset of the standard that fits game-based educational contents. Special attention has been given to those categories we considered more relevant in this context: Educational metadata and technical characteristics. Table 1 summarizes the different categories from the IEEE LOM standard directly supported inside the editor.

<table>
<thead>
<tr>
<th>Category</th>
<th>Element</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General</td>
<td>1.2. Title</td>
<td>Name given to this LO</td>
</tr>
<tr>
<td></td>
<td>1.3. Language</td>
<td>The primary human language of the content (e.g. en, EN)</td>
</tr>
<tr>
<td></td>
<td>1.4. Description</td>
<td>A textual description of the content of this LO</td>
</tr>
<tr>
<td></td>
<td>1.5. Keyword</td>
<td>A keyword or phrase describing the topic</td>
</tr>
<tr>
<td>2. Life Cycle</td>
<td>2.1. Version</td>
<td>The version of this LO</td>
</tr>
<tr>
<td>4. Technical</td>
<td>4.4. Requirement</td>
<td>Technical capabilities necessary for using this LO</td>
</tr>
<tr>
<td></td>
<td>4.4.1.3. Min Version</td>
<td>Lowest version of the &lt;e-Adventure&gt; engine required</td>
</tr>
<tr>
<td></td>
<td>4.4.1.4. Max. Version</td>
<td>Highest version of the &lt;e-Adventure&gt; engine supported</td>
</tr>
<tr>
<td>5. Educational</td>
<td>5.1. Interactivity Type</td>
<td>Predominant learning mode (active, expositive or mixed)</td>
</tr>
<tr>
<td></td>
<td>5.2. Learning Resource Type</td>
<td>Specific kind of LO (exercise, simulation, exam, self-assessment or lecture)</td>
</tr>
<tr>
<td></td>
<td>5.3. Interactivity Level</td>
<td>The degree of interactivity characterizing this LO (very low, low, medium, high or very high)</td>
</tr>
<tr>
<td></td>
<td>5.4. Semantic Density</td>
<td>The degree of conciseness of the LO (very low, low, medium, high or very high)</td>
</tr>
<tr>
<td></td>
<td>5.5. Intended End User Role</td>
<td>Principal user(s) for which this LO was designed (teacher, author, learner or manager)</td>
</tr>
<tr>
<td></td>
<td>5.6. Context</td>
<td>The principal environment of application (school, higher education, training or other)</td>
</tr>
<tr>
<td></td>
<td>5.7. Typical Age Range</td>
<td>Age of the typical intended user.</td>
</tr>
<tr>
<td></td>
<td>5.8. Difficulty</td>
<td>How hard it is to work with or through this LO (very easy, easy, medium, difficult or very difficult)</td>
</tr>
<tr>
<td></td>
<td>5.9. Typical Learning Time</td>
<td>Approximate or typical time it takes to work with or through this LO (duration in hours and minutes)</td>
</tr>
<tr>
<td></td>
<td>5.10. Description</td>
<td>Comments on how this LO is to be Used</td>
</tr>
<tr>
<td></td>
<td>5.11. Language</td>
<td>The primary human language of the content (set with 1.3)</td>
</tr>
</tbody>
</table>
In order to simplify as much as possible the work of the content author, the metadata annotation process is done in a semi-automatic way. The <e-Adventure> editor harvests as much as possible metadata from the game description simplifying the instructor’s task. All the technical metadata are also automatically generated, and the instructor can focus on the definition of metadata related to educational aspects.

**Figure 1.** Screenshot of the annotation with IEEE LOM metadata in <e-Adventure>

**Figure 2.** Screenshot of an <e-Adventure> LO package loaded in the Reload Editor application, displaying part of its metadata section
The exportation process of the game creates a minimal self-contained LO. The instructor uses the editor to create an IMS CP bundle that includes: the game itself (including assets), its metadata, a streamlined version of the <e-Adventure> game engine, a simple html page used to launch the game and a manifest file used to describe the bundle content. The generated LO contains a single course with a single educational resource (the game), however the LO may be refined or integrated with larger courses using tools that comply with the selected specifications. For example, as seen in Figure 2, the LO can be edited using the Reload Editor (http://www.reload.ac.uk), which allows its integration with bigger pre-existing courses and or refinement of the original metadata included by the <e-Adventure> platform.

In summary, with <e-Adventure> it is possible to package the games according to the IMS CP specification and annotated with meta-data following the IEEE LOM standard. The result is a shareable and reusable educational videogame packed as an LO and ready to be deployed in any IMS-compliant LMS. In the next section, we analyze the actual level of compliancy and how this approach facilitates the deployment and delivery process for educational games.

Case Study

In order to test the compliancy of the proposed solution in terms of content interoperability, we have conducted a local case study focused on the problem of game delivery through different LMSs. Additionally, the possibility of having a connection established between the games and the LMS and using that communication channel to exchange tracking and assessment information has also been evaluated. The objectives were thus the following:

1. To test the successful integration of a real <e-Adventure> game in a broad sample of LMS as a) a single-content course and b) a complex course combining both game content and traditional (i.e. web-based) content.
2. To test the compliancy of the automatic assessment engine in the same sample of LMS (i.e. determine which LMS could successfully process the statement of results produced by the game and attach it to the profile of the student to be used in the future).

On the other hand, it was required to specify selection criteria to determine which LMS would be selected to take part in the test in a manner that the results obtained could be significant. Covering every single available LMS was deemed impossible, as new tools are emerging constantly. We made a selection covering some widely extended open source tools and complemented it with commercial products such as the two versions of BlackBoard’s WebCT currently being used in our university: WebCT versions 4 and 6. We also tried to make special emphasis on tools that currently support communication with interactive LO using either the ADL SCORM reference model or the IMS Learning Design specification. The complete list of evaluated products can be found in Table 2.

For each individual LMS we performed three different tests: The first test was the deployment of an <e-Adventure> game embedded in a web document without any attention to specific standards. In order to reflect the usual skill set of instructors and, in most cases, their actual access rights, this integration should be possible using the LMS interface, instead of requiring an FTP connection to upload the contents or any form of server-side modification. The second test was the deployment of the game as a standardized LO packaged according to the IMS Content Packaging specification. Finally, we also reported on the possibility of establishing a communication link between the game and the LMS, using any of the two communication protocols supported by the <e-Adventure> platform.

Presentation and discussion of the results obtained

The next table (Table 2) summarizes the results obtained in all the tests performed. For each LMS the table depicts if it was possible to deploy the LO as a single web document (column 2), into a structured course according to the IMS CP specification (column 3) and if the automatic assessment mechanism worked properly (column 4).

As the table shows, the standardized LO could be successfully deployed in all the LMS, but in some cases the integration of the game “as is” was not possible without low-level access to the server given the way in which some systems treat the inclusion of files and/or downloads (CopperCore/SLeD is a special case, given that by design they do not support content that is not part of a standardized Unit of Learning). This highlights the importance of packaging the games according to the current standards and specifications. Both CopperCore (with the SLeD front-
end) and .LRN (with its IMS-LD module (Escobedo del Cid et al. 2007)) can establish a communication link using the proprietary protocol defined in <e-Adventure> and special plug-ins. With LAMS it is also possible to establish a proprietary communication link.

Table 2. Summary of the results obtained in the integration and deployment process of the LO-game in the diverse LMS

<table>
<thead>
<tr>
<th>Environment</th>
<th>Web document</th>
<th>Standard LO</th>
<th>Tracking support</th>
</tr>
</thead>
<tbody>
<tr>
<td>CopperCore / SLeD</td>
<td>N/A</td>
<td>Yes</td>
<td>Proprietary protocol</td>
</tr>
<tr>
<td>LAMS</td>
<td>Yes</td>
<td>Yes</td>
<td>Proprietary protocol</td>
</tr>
<tr>
<td>.LRN</td>
<td>No</td>
<td>Yes</td>
<td>Proprietary protocol</td>
</tr>
<tr>
<td>Moodle</td>
<td>Yes</td>
<td>Yes</td>
<td>Proprietary protocol</td>
</tr>
<tr>
<td>Sakai</td>
<td>Yes</td>
<td>Yes</td>
<td>Experimental SCORM 2004 support</td>
</tr>
<tr>
<td>WebCT 4</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>WebCT 6</td>
<td>No</td>
<td>Yes</td>
<td>SCORM 1.2</td>
</tr>
</tbody>
</table>

Applications

As described in the previous subsection, the main objective of the experiment was to test the integration of <e-Adventure> games in educative contexts through their deployment in LMS as LOs. The available systems are in constant evolution, and so are the needs of the application environments. It is not uncommon to see an institution that changes the technology of their LMS, for example alternating between proprietary and open source tools.

Having observed poor performance during an evacuation drill, the Dean’s Office at the School of Computer demanded an investigation to identify and improve all the issues identified during the drill and the implementation of a training program to ensure both staff and students know the evacuation protocol. However this is not an issue easy to address. Memos about how to proceed in these evacuations have been proved useless in the past because most people simply do not take the time to read the instructions, and fire safety briefings for personnel or students register low attendance rates. For that reason we developed a low-cost game-like simulation of the evacuation procedure using <e-Adventure> (Moreno-Ger et al. 2008a) as a mechanism to capture both staff and student attention.

Figure 4. Screenshots of the evacuation procedure game-like simulation implemented with the <e-Adventure> platform
This case exemplifies the problem of how to deliver educational videogames to the students that has been described all along the paper. The high number of users of the game difficulties its delivery through a traditional delivery channel (e.g. through CD-ROMs). To tackle this issue, the game is to be delivered through the LMS currently available at Complutense University, which is based on WebCT 4. All the teachers, staff and students have access to this system, so it could be a good controlled distribution channel for the game.

However, the university is currently considering alternative systems to cover the online support for the courses. Apart from a potential upgrade of the current WebCT platform, different open source tools are also being considered (such as Moodle and SAKAI). As the results of our test show, the investment in the development of the game is safe regardless of the final decision, as the game can be seamlessly integrated with any of the environments currently being considered.

Conclusions

So far, the research community broadly accepts the potential benefits of educational videogames, even when such benefits have not been properly and empirically demonstrated yet. However, the actual application of educational videogames is limited due to a number of factors, including the difficulty of integrating the games into ongoing educational processes (McFarlane et al. 2002; Kirriemur and McFarlane 2004). Our proposal is to take advantage of the current LMS infrastructure and their instructor-oriented content deployment tools, following the Learning Objects Model supported by standards and specifications.

As the case study proves, using the <e-Adventure> platform for the creation of educational point-and-click adventure games packaged as IMS-CP Learning Objects instructors can successfully deploy the games in several of the LMS with more extended use. The packaging effort is very low (just select the appropriate option on the <e-Adventure> editor) and the deployment is straightforward using the content importation facilities of most LMS. Therefore the games can be easily delivered to the students without requiring additional installation and deployment efforts as it was the case in some of the examples mentioned in the related work section.

Additionally, in some environments it is possible to use the communication between the LMS and the game engine to produce an automatic evaluation of the activity within the game and attach it to the profile of the student. However the results of the case study suggest that the compatibility degree with ADL SCORM must be taken into account when considering this approach. Most systems offer different degrees of compatibility with this reference model, although most systems only comply with SCORM 1.2. In terms of tracking and assessment, this means that the LMS can gather data submitted from the games, but cannot use it to affect sequencing decisions. In contrast, the environments that use the proprietary communication protocol or SCORM 2004 can get more value from this information.

In addition, the approach also facilitates the integration of the games with other LOs containing web-based content (using the deployment tools of LMS or standards-compliant content authoring tools like the Reload Editor). This could be used to create mixed courses (combining games and other content) which has an important application potential, for instance to provide alternative paths to students that would not benefit from a pure game-based learning approach (Torrente et al. 2008a). However, an interesting alternative not explored in this study is how to connect diverse games in the same course, in a manner that the results of a game could affect the behaviour of the next one, thus supporting the creation of complex courses as game-based trees. A priori, such connection between games could be achieved acting the LMS as intermediary by using the assessment engine and the adaptation mechanism described in (Moreno-Ger et al. 2007). Nonetheless this has not been empirically tested yet and further research should be conducted.

The proposed approach does not force educational institutions to install any specific software, as it takes advantage of the existing LMS infrastructure. In addition, this approach promotes reusability of contents, protecting the investment against possible LMS replacements. On the other hand, the educational videogames produced can be stored in public LO repositories to promote the exchange of contents. We are thus inheriting all the benefits of the LO model.
Next steps in the project are to test the educational outcome of using games and simulations in an actual University course. We are currently collaborating with researchers from the Complutense School of Medicine to create educational games to support the practical exercises performed by the first and second year students. The games will be delivered to the students through the Complutense’s e-Learning environment (Virtual Campus), integrated with the rest of the online content that supports the corresponding subjects.

Acknowledgements

The Spanish Committee of Science and Technology (projects Flexo-TSI-020301-2008-19 and TIN2007-68125-C02-01) has partially supported this work, as well as the Complutense University of Madrid (research group 921340,) and the EU Alfa project CID (II-0511-A). Special thanks to Ernie Ghiglione for his work testing the integration of the <e-Adventure> platform on LAMS environments and to Dr. Carl R. Blesius for his tests with the .LRN platform. This work was also greatly improved thanks to the insightful comments provided by the anonymous reviewers.

References


Influence of Structure and Interaction on Student Achievement and Satisfaction in Web-Based Distance Learning

Hye-Jung Lee¹ and Ilju Rha²

¹Center for Teaching and Learning, Seoul National University, Korea // Tel: +82 2 880 4023 // hyejung@snu.ac.kr
²Department of Education, College of Education, Seoul National University, Korea // Tel: +82 2 880 7642 // iljurha@snu.ac.kr

ABSTRACT
This study examines the influence of instructional design and management style on student achievement and satisfaction in a web-based distance learning environment. From the literature review, two major instructional design and management styles in web-based distance education were conceptualized as structure and interpersonal interaction. To investigate the differences on learning output variables, two web-based instructional programs were developed as college level courses. One course was developed and implemented mainly with a resource-based highly structured self-learning mode with little interpersonal interaction, and the other course was developed and conducted mainly with interpersonal interaction without well-structured materials. Sixty seven and fifty six juniors were enrolled in each course at a cyber university for one semester. Questionnaires, achievement scores, satisfaction levels, and interview data were collected and analyzed. Findings indicated that self-learning with well-structured materials was not inferior to highly interactive instruction without well-structured materials in receptive learning achievement. This implicates that well-structured material can possibly replace a teacher’s one to one interaction in receptive learning achievement. However, students in the interactive course without well-structured materials, rather than in the well-structured course with little interaction, showed higher achievement in critical thinking learning. In terms of satisfaction, students in the structured course were more satisfied with structure, and students in the interactive course were more satisfied with interpersonal interaction. Implications of these findings and recommendations for further study are suggested.

Keywords
Structure, Interaction, Distance learning, Instructional design, Instructional management, Web-based learning

Introduction
Recent advancement in the web and multimedia technology is changing the way of teaching and learning. Because of the technologies’ flexibility and diversity, distance learning programs using internet and multimedia contents also have proliferated (Menchaca & Bekele, 2008). In web-based distance instruction, one of the critical issues is how to design, develop and manage instructional programs to get effective, efficient learning that is attractive. For thousands of years, learning and teaching always took place in close proximity, and this has become firmly anchored in human consciousness. Learning and teaching at a distance is, therefore, regarded as something extraordinary and a kind of defect. Educators have tried to make every effort to get over this spatial separation. The first pedagogic approaches specific to distance education aimed at finding ways by which spatial distance could be bridged, reduced or even eliminated. The question was asked: what must be done to make distance equivalent to proximity in distance education? The pedagogics of distance education is derived basically from the efforts to answer to this question.

One answer to this question is making and managing the best learning materials as possible, which require little to no teacher’s interaction. Educators at traditional distance universities have put forth the most effort and funds in the professional development and production of qualitatively excellent teaching materials for the purposes of self-study, which were distributed by post and then by web-based cyber instruction programs these days (Peters, 1998). They have tried to include the effect of teacher’s interaction into well-structured course and material for self-learning. Moore (1993) conceptualized this as structure, because he regarded the main feature of making and managing programs as the structuring of the learning and teaching. Also many studies have presented elements and rationale for course/contents structure as well as instructional strategies for structuralization (Bischoff et al., 1996; Carr-Chellman & Duchastel, 2000; Chen, 2001a, 2001b; Chen & Willits, 1999; Lee, 2004a; Ostlund, 2008; Peters, 1998; Saba, 2005; Saba & Shearer, 1994). Researchers emphasizing structure in distance learning seem to believe that well-structured material can overcome the absence of teachers.

ISSN 1436-4522 (online) and 1176-3647 (print). © International Forum of Educational Technology & Society (IFETS). The authors and the forum jointly retain the copyright of the articles. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the full citation on the first page. Copyrights for components of this work owned by others than IFETS must be honoured. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from the editors at kinshuk@ieee.org.
Others say, however, that it is not sufficient simply to enable students to study in isolation with the help of distance learning materials. Students must be able to discuss with their teachers and other students in order to learn, because this is the real foundation of academic teaching (Daniel and Marquis, 1979; Moore & Thompson, 1990; Morris, Mitchell, & Bell, 1999; Sewart, 1982; Verduin & Clark, 1989). They say that a person’s grasp of the acquired and assimilated knowledge can only develop in discussion through interpersonal interaction (Garrison, 1993). But interpersonal interactions at a distance have not been easily achieved until the advent of the computer and network technology. Now educators’ and developers’ attention has turned to converting distance learning into a more interactive one than ever as computer and web technology become widely available. As a communication medium, the web discussion board presents many possibilities to effective interaction, which is not encountered in traditional instructional environments or in print publications. So many studies on interaction in web-based learning persist positive pedagogical effects of interaction and present various interaction strategies for better learning (Beuchot & Bullen, 2005; Dennen et al., 2007; Garrison & Cleveland-Innes, 2005; Kehrwald, 2008; Novitzki, 2005; Russo & Campbell, 2004; Tu & McIsaac, 2002; Weaver, 2008).

These two perspectives, structure and interaction, bring educational policy and priorities into conflict with each other; they are characteristics of distance education on the whole and are the source of many disputes as Lee (2004b) indicated. Moore (1993) has been theorized the relationship of these two perspectives. According to him, the more we structuralize a course and its materials, the less we need interpersonal interaction. When materials are not structured well enough, teachers need to provide more interaction for compensation. So the decision making on the extent – whether to focus on structuring course and materials or making interactive course, or how to combine each of the extreme – is critical, especially in distance education organization, as each (structure and interpersonal interaction) need a substantial budget to implement in a satisfactory manner. If the educational organization decides to focus on structuralization of programs, they need to hire expert contents designers and web programmers. If the organization wants to emphasize interpersonal interaction, they must have enough instructors or tutors to interact with the students. These two concepts are, in effect, contrary in relationship. According to Peters (1998), an interesting task for distance-teaching pedagogics might be to undergo a theoretical and experimental work to find whether and if so where the pedagogic optimum can be found between two extremes, with an extremely structured course at one end and extremely interactive course at the other end.

Unfortunately, however, studies about the rationale of each program’s effect or the influence of the design and management style on learning output variables cannot be found easily. Although the majority of the literature on web-based distance education focuses on the effectiveness of web-based distance education or educational media, and emphasizes structure (converting existing, traditional course materials to a web-based format) or a teacher’s interaction with students respectively, it doesn’t reflect a scientific approach necessary to discern which – and why – is more effective, how learning outputs are different in each program, and whether – if so, how much – structure can be a substitute for interpersonal interaction.

Therefore, this research purposes to analyze and compare learning outputs, such as student achievement and satisfaction, in highly structured instruction with the least interaction and highly interactive instruction with the least structure, and to discuss two macroscopic instructional design and management styles and their influence. For this research, two different WBI (Web-Based Instruction) programs with different design and management styles, one with highly structured instruction and the other with highly interactive instruction, were developed. Quantitative and qualitative data were collected and analyzed to investigate the influence of design and management styles on learning output in web-based distance learning.

**Theoretical Background**

**Structure and interaction**

There is a physical separation of a teacher and a learner in distance education. One of the efforts in reducing the distance between teacher and students was taken when people started to interpret distance education as the simulation of a conversation between teacher and student. Distance educators recommend that the contents be displayed in the form of a written dialogue. When writing teaching text, teachers must imagine that they are speaking to someone, and this is supposed to make them use a natural spoken language in writing the text wherever possible (Holmberg, 1983). Many distance educators emphasize the importance of structure in developing text materials in distance education to make up the distance. Kearsley & Lynch (1996) contended that distance education courses
must have high structure to be successful, and William (2006) said that the more ID components incorporated into a distance education course, the greater the student achievement outcome. It is also reported that most critical factor in distance learning is structure and interaction, neither learners’ characteristics such as their technical expertise nor course delivery format (Stein et al., 2005).

The term structure, defined by Moore (1993), is in contrast to dialogical learning. Structure refers to elements of the course’s design, such as learning objectives, activities, assignments, evaluation, whereas interaction is the communication between the instructor and learners (Moore & Kearsley, 2005). Structuring of instruction is consistently planned on a targeted basis in small steps; its time is regimented and it is uniformly controlled and evaluated. The main instrument of this learning and teaching are printed materials or multimedia learning packages which contain carefully developed and optimized courses and set learning into motion. Moore (1993) used the example of a teaching film for television to show just how far this structuring could go: in the film literally every word, every action of the teacher, every minute of the available time and even the tiniest detail of the contents were laid down beforehand.

The structuring of learning and teaching is in fact nothing new and certainly not unique to distance education. Presenting knowledge through books has usually been broken down by means of components such as Foreword, Introduction, Chapters, Summary and Conclusion. And in the last two centuries, the articulation of instruction has been a central theme of pedagogics, especially in the field of instructional design or educational technology. This structure is seen most clearly and most densely in programmed instruction. It is the representative of instructional theories in which each word, each learning step, each teaching strategy is planned and developed with the greatest care in order to simplify learning and make teaching more successful. What is nevertheless new in Moore’s suggested concept is not structuring in itself but the extreme extent of its application (Peters, 1998).

This structure discussed here, is about the elements in the course design and contents design. A course consists of such elements as learning objectives, content themes, information presentations, case studies, pictorial and other illustrations, exercises, projects and tests. All these should be very carefully composed and structured. In addition, we need to consider another type of structure especially in a web-based environment. That is, interface structure, which is not covered in Moore (1993)’s concept; he seems to consider the concept, structure, in a correspondence learning environment, not an e-learning context in recent days. Interface structure encompasses site structure, navigation structure, and screen design structure on the computer screen that students face directly (Lee, 2004b). Therefore, highly structured instruction in this research means that the extent of structuring is quite high in terms of course, contents, and interface structure for self learning with little to no teacher’s interaction.

While structure is an issue about program design and development, interaction is an issue rather about implementation and management. Moore (1993) categorized interaction types into student-teacher, student-student, and student-contents interaction. Moore called student-teacher interaction dialogue, which was in contrast to his definition of structure. The concept of Moore’s dialogue does not mean the written presentation of simulated letters or conversations, rather it means direct and indirect oral interaction between teachers and students, that is, an actual dialogue. The concept in the definition of dialogue that is in contrast to structure, however, needs to be redefined in this context because student-student interaction is more like student-teacher interaction, and it is essentially different from student-contents interaction which is self-learning with materials. So these two interactions (student-student, student-teacher) could be redefined as interpersonal interaction, which is beyond just dialogue (Lee, 2004b). Interpersonal interaction is certainly differentiated from intra-personal interaction of student-contents interaction that is rather included to structure because structuring materials is basically for effective intra-personal interaction between a student and contents. Meanwhile, we should consider electronic messages on the web board in written text as well as oral dialogue in web-based distance environment. This could be very attractive because it would mean that more personal communication and academic discourse – something that lacks in traditional distance education – might be combined with recent technology-based distance learning. So the interpersonal interaction in this research includes student-teacher and student-student interaction, mainly reflected the cyber space through web board or e-mails as posted messages in written text.

Interaction has been the most frequent research theme in web-based distance learning because of its pedagogic genesis. Many studies on interaction in web-based learning persist positive pedagogical effects of interaction and present various interaction strategies for better learning (Beuchot & Bullen, 2005; Dennen et al., 2007; Fulford, & Zhang, 1993; Garrison & Cleveland-Innes, 2005; Gunawardena & Zittle, 1997; Russo & Campbell, 2004; Tu & McIsaac, 2002). They persist no matter how the program is structuralized, learning cannot be completed without
interpersonal interaction. Moore & Thompson (1990) and Verduin & Clark (1989) found that distance education courses were most effective when student-student interactions were present and the instructor feedback was timely.

However, the belief of the more, the better in interaction is becoming to be criticized. It has been found that, although instructor-learner interaction is important, not all students are looking for high levels of interaction, such as a response to every discussion board posting (Su et al., 2005). Thus, instructor presence does not follow a more is better pattern. Various studies have shown the cases for it. Mazzolini & Madison (2003) found that increased instructors posting did not result in increased student participation. In fact, as the instructor became more involved, student messages became shorter and more infrequent. Dennen (2005) noted similar interactions in a course where the instructors wrote about one-half of all messages in the class in an attempt to respond to each individual student. Dennen et al. (2007) suggested that there is a threshold at which an instructor’s heavy-handed or overwhelming amount of communication inhibits or discourages learner communication and participation. Besides, it is reported that some instructors may not know how to fluidly, efficiently, and effectively provide important information in an online setting. So structuring those elements is necessary to help instructors. Structuralization would make up for instructor’s unpreparedness (Dennen et al., 2007).

<table>
<thead>
<tr>
<th>Table 1. Structuring elements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
</tr>
</tbody>
</table>
| Course structure | Dosing  
Providing Course orientation: welcoming, characterizing the course contents, description of the required preliminary knowledge, indications of correlations to other areas of knowledge and possibly to job practice, description of rough learning and teaching objectives, naming advantages obtained by reaching and learning goals, etc.  
Scheduling  
Self-tests  
Peer evaluation  
Assignments and pre-scheduled deadline  
Systematic management of course pace and learning speed  
Strict grading according to notice announced in advance |
| Contents structure | Portioning  
Providing contents, lists  
Indication of learning and teaching aims  
Sequencing  
Stressing what is important  
Showing perspectives for further study  
Branching  
Providing cross-references  
Arrangement in blocks and characterization by marginal notes  
Providing summative map, structural site map, procedural map  
Providing concept map, concepts tree  
Providing comparison/contradiction  
Providing abstracts  
Providing index  
Providing various Instructional methods: questions for repetition, questions for training, simulated dialogue  
Writing contents with spoken language, like conversation style |
| Interface structure | Providing spatial layout to avoid cognitive overloading  
Providing search index function  
Emphasizing by spaced printing  
Providing marginal notes  
Visualization with graph, table, or figures  
Using frames, underlining, bold, italic print, decimal numbering  
Using comic characters like Avatar for students not to feel isolated |

*Note. From previous studies (Bischoff et al., 1998; Chen & Willits, 1999; Lee, 2004a; Peters, 1998; Saba & Shearer, 1994)*
Structure and interaction seem to be standing contrary on one hand, complementary on the other hand. According to Moore (1993), structure is a variable that determines to some degree the extent of dialogue (Gorsky & Caspi, 2005). Saba & Shearer (1994) assumed a systemic and dynamic relationship between structure and dialogue and observed that transactional distance varies by the rate of structure and dialogue. They found that as dialogue increases in a videoconferencing environment, structure decreases to keep the system stable. When structure increases, dialogue decreases. Moore & Kearsley (2005) noted that success in distance teaching is determined by the extent to which the instructor and the institution are able to provide appropriate structure and the appropriate quantity and quality of dialogue between instructor and learner. But the challenge here is that what is appropriate is left to the instructor to determine (Stein et al., 2005). Moore acknowledged that there are no ready-made answers to how much dialogue and structure is needed and desirable for effective learning. Although he recommends that instructors in doubt should err on the side of too much structure rather than too little (Moore, 2004), it is still challenging to find pedagogical optimum between two extremes. And before finding the optimum, a preliminary research on how these two extremes influence on learning respectively, how learning is differently influenced by each centered instruction, is necessarily to be studied.

**Structuring elements**

Well-structured materials in this research are the learning contents which are designed and developed considering various structural elements (see Table 1), which were from previous literature (Bischoff et al., 1998; Chen & Willits, 1999; Lee, 2004a; Peters, 1998; Saba & Shearer, 1994). Non-structured materials are the learning contents which were not specifically considered those structural elements in design and development.

**Research Method**

**Program development**

In order to compare learning output variables in web-based instruction, two WBI (Web-based Instruction) programs were developed as college level courses. One course (called structured instruction in this paper) was developed and implemented mainly as a highly structured resource-based self-learning mode, and the other course (called interactive instruction in this paper) was developed and implemented to carry mainly interpersonal interaction with the least course/contents/interface structure. Elements and design/management factors of each instructional program, such as elements for course/contents/interface structuralization, or management tips for the interactive course, were shown in Table 1. Programs were designed and developed according to traditional ISD (Instructional Systems Development) model, considering these elements. Three WBI experts and practitioners verified these two different programs throughout the development process. Subject matters of both courses were same, which were about general social science. The development process for two instructional programs is as follows; Analysis and Planning (Learning contents, students, learning environments were analyzed first and research design was planned considering these analysis results) → First instructional design and draft storyboarding → Validation by three experts (Ph.Ds in the field) and developers → Second instructional design and storyboarding → Production of digital materials → Pilot implementation of course → Revision and redevelopment of program → Validation by three experts and completion of final instructional programs (The programs were also validated by students during the process of implementation and data collection; see Results).

**Participants and setting**

Sixty seven and fifty six juniors at a cyber university in Korea, ages 20s to 50s, were randomly assigned and required to take each course respectively for one semester. There was a pre-test to know prior knowledge and the results showed no significant difference between two groups (Table 2). Same instructor and same contents were applied to both classes. In terms of student characteristics that might have influence on the results of this research, perceived learner characteristics were analyzed a week prior to each program’s implementation. Students used a five-point Likert scale to analyze whether students were extrovert vs introvert, social (or outgoing) vs rather individual, active vs passive in class. As a result, students in the highly structured instruction were more extroverted (p<0.05) and more active in class (p<0.01) than the students in the interactive instruction were. However, the student factor doesn’t
seem to be really influential in this context because students’ characteristics were reversed in the previous pilot implementation and showed same results as this paper. That is, in the pilot implementation, students in the interactive instruction were more extroverted (p<0.05) and more active in class (p<0.05) than the students in the structured instruction were. But the results of the pilot implementation showed no significant difference from the result in this paper. Therefore, student’s characteristics weren’t a critical factor as much as instructional method in this research.

### Table 2. Pre-test difference for students’ prior knowledge between structured instruction and interactive instruction

<table>
<thead>
<tr>
<th>Record</th>
<th>Class</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>df</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall record</td>
<td>Structured instruction</td>
<td>53</td>
<td>3.83</td>
<td>0.91</td>
<td>93</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>Interactive instruction</td>
<td>42</td>
<td>3.50</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Implementation

Two classes were homogeneous in terms of instructor and subject matter, since same instructor and same subject matter was assigned to both classes. Main difference was instructional design and management way (Table 3).

### Table 3. Examples of course structure difference between the structured course and the interactive course

<table>
<thead>
<tr>
<th>Course structuring elements</th>
<th>Structured instruction</th>
<th>Interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dosing</td>
<td>Highly structured contents were provided regularly every week</td>
<td>Most of the instruction is based on the interaction on the online board. Non-structured contents were provided irregularly.</td>
</tr>
<tr>
<td>Orientation</td>
<td>Implemented in detail</td>
<td>Very brief orientation provided</td>
</tr>
<tr>
<td>Grading</td>
<td>Strict grading according to notice announced in advance</td>
<td>Participation and peer evaluation included. (But these were not included in the result for comparison with structured instruction.)</td>
</tr>
<tr>
<td>Assignments</td>
<td>Assignments required every week. No feedback provided.</td>
<td>Immediate feedback provided to every assignments.</td>
</tr>
<tr>
<td>Deadline</td>
<td>Pre-scheduled deadline</td>
<td>Flexible deadline according to students’ consensus</td>
</tr>
<tr>
<td>Course schedule</td>
<td>Systematic management of course pace and learning speed</td>
<td>Flexible management applied in regard of pace, speed, sequence, etc.</td>
</tr>
</tbody>
</table>

In the highly structured instruction, students were required to study highly structured web-based material and to undertake assignments given every week, usually alone. Instructor’s feedback was provided as little as possible. Instead, most of the possible feedbacks by the instructor were structured into the web materials. For example, materials were written with natural oral language just like a simulation of conversation. And a teaching assistant cartoon character was used all along the learning process in the material. In addition, various structuring elements in Table 1 were reflected to the program and the web-based material. The instructor was trying to neither encourage nor discourage interactions among students. Meanwhile, in the highly interactive instruction, students were required to undertake assignments and to participate in team discussions assigned throughout the course. The instructor’s immediate feedback to assignments and questions were provided, although well-structured materials were not provided. All interactions were recorded on an asynchronous web bulletin board. Table 3 shows some examples of course structure difference between two classes.

### Procedure

This research was done over a period between 2002 and 2004 from program development to data collection and analysis. Design and development of two different types of instructional programs were started from early 2002. The treatment was first operated with these programs as a pilot implementation at fall semester of 2002. After some revision, finally developed instructional programs were implemented in 2003 again. And then interview and analysis
of collected data was conducted until 2004. The results were found to be credible enough to be replicable since there was no difference between two implementations.

The students’ cognitive achievement data was collected in two different ways. One was about receptive learning, which consisted of declarative knowledge, information, concepts, or theories that were mostly accepted without criticism; ex. Introducing concepts and definitions of the terminology related in the class. In receptive learning, students don’t have to criticize what they learn, what they do is just understanding, memorizing, and recollecting. That is, reception in itself. The other was about critical thinking learning, which contained controversial issues requiring arguments, criticism and discussion; ex. Does distance education using high technology reproduce classes of society? Does it make the rich-get-richer and the poor-get-poorer?, or Is distance education cost-effective?

For evaluation reliability, three evaluators’ gradings were correlated (Pearson r = 0.84**, p< 0.01). In terms of satisfaction level, a satisfaction measurement tool which was developed and validated by Kim and Ryu (2000) was used after modification. Twenty students were interviewed at the end of each course to verify all quantitative data and to provide more detailed information regarding the factors of learning outputs.

Besides, questionnaires, achievement scores, satisfaction levels, online messages, interview data were collected and analyzed. Various statistical analysis methods such as correlation analysis, contents analysis, t-test, and frequency analysis, were applied to the data.

**Results and Interpretations**

**Cognitive achievement**

Students’ cognitive achievement, shown in Table 4, was evaluated in terms of receptive learning and critical thinking learning. In cognitive achievement, the well-structured course was not inferior to the teacher’s interaction in receptive learning. However, in critical thinking learning, students in interactive instruction showed higher achievement than those in structured instruction. In receptive learning, well-structured instructional program was provided for the structured course students, while simple text scripts and instructors’ feedback was provided for the interactive course students. In critical thinking learning, a well-structured contents, including pros/cons on issues, was provided for the structured course students, while live discussions and arguments (mainly student-student interaction, instead of instructor-student interaction) occurred on the board in the interactive class.

<table>
<thead>
<tr>
<th>Table 4. Student achievement differences between structured instruction and interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Overall record</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Record in critical thinking learning</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Record in receptive learning</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*p< 0.05, **p< 0.01

Two groups’ answers about influential factors on cognitive achievement were also different. Table 5 shows that students in the structured instruction considered web-based material as the most critical factor for cognitive learning, and assignment was considered as a subsequent critical factor. Meanwhile, students in the interactive class consider messages on the discussion board as the most critical achievement factor. Interviews conducted at the end of the semester also verified these findings;

“...actually, studying is individual thing after all, isn’t it? So the most critical thing in this kind of e-learning is course and contents structure. Interpersonal interaction is just additional, not essential, I think...”

- From an interview with a student in the structured instruction.
“...the discussion in my class was fabulous! I think this kind of discussion and interaction is really necessary in distance learning. Otherwise, it would be hard to keep studying, not in face-to-face environment...”

- From an interview with a student in the interactive instruction.

**Table 5.** Frequency analysis results of students’ perceived achievement factors

<table>
<thead>
<tr>
<th>Perceived student achievement factors</th>
<th>Structured instruction</th>
<th>Interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Web-based material (67%)</td>
<td>1. Messages on class web board (77%)</td>
<td></td>
</tr>
<tr>
<td>2. Assignments (53%)</td>
<td>2. Assignments (51%)</td>
<td></td>
</tr>
</tbody>
</table>

**Satisfaction**

A 26-item survey was used to measure participants’ satisfaction toward their instructional program. This survey, modified from Kim and Ryu (2000)’s tool, consisted of three categories: structure (course, contents, interface), interpersonal interaction (instructor-student, student-student), and overall attractiveness of instruction. Respondents used a five-point Likert scale (5=very satisfied, 1=very dissatisfied) to rate their satisfaction level. The alpha reliability coefficient of the survey was .93. Table 6 shows t-test results of those students in the structured instruction who were satisfied more with the structure factor of the program whereas the students in the interactive instruction were satisfied more with interpersonal interaction factor of the program. There was no significant difference about the attractiveness of the program. This result was verified in Table 6.

**Table 6.** Students’ satisfaction level differences between structured instruction and interactive instruction

<table>
<thead>
<tr>
<th>Satisfaction level with structure</th>
<th>Structured instruction</th>
<th>Interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Satisfaction level with structure</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Satisfaction level with interaction</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Satisfaction level with attractiveness of instruction</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
</tbody>
</table>

Additionally students were asked to identify their perception of critical satisfaction factors by checking the most critical factor and the second critical factor. As a result, students in the structured instruction considered the design of web-based material as the most critical satisfaction factor. On the other hand, students in the interactive instruction answered instructor’s feedback as the most critical satisfaction factor. Both of them considered the contents of the program as a secondary satisfaction factor (see Table 7). These findings implicate that instructional programs in this research were developed appropriately, along with design intention. Students were satisfied with the program more than at an average level, so an error from program dissatisfaction was avoided. Table 8 also supports two instructional programs were appropriately developed, since the perception of students in the structured course showed their course is highly structured while the perception of the students in the interactive course showed their course is highly interactive one. The instructional programs developed in this research were validated again by these results.

**Table 7.** Students’ perceived satisfaction factors in structured instruction and interactive instruction

<table>
<thead>
<tr>
<th>Satisfaction factors</th>
<th>Structured instruction</th>
<th>Interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Satisfaction factors</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Satisfaction factors</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
</tbody>
</table>

**Table 8.** Students’ perception of class between structured instruction and interactive instruction

<table>
<thead>
<tr>
<th>Perception items</th>
<th>Structured instruction</th>
<th>Interactive instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>Contents material is well-designed enough for me to study well.</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
<tr>
<td>I can’t understand easily when I study the</td>
<td>Structured instruction</td>
<td>Interactive instruction</td>
</tr>
<tr>
<td>Class N Mean St. Dev. df T</td>
<td>Class N Mean St. Dev. df T</td>
<td></td>
</tr>
</tbody>
</table>

[p<0.05]

379
Discussion and Conclusion

The purpose of this study is to analyze the influence of instructional design and management style on student achievement and satisfaction in a web-based distance learning environment. Results indicate that program structure is possibly able to somewhat replace interpersonal interaction as early distance educators expected, especially when the contents have receptive characteristics like understanding, memorizing, and recollecting facts/concepts/principles without criticism. This is the contradictory empirical evidence to the general belief that the essential of learning is only from interpersonal interaction between teacher and student whatever the learning contents are. This research suggests that a well-structured instructional program can be provided as a substitute for teacher’s interaction in receptive learning.

On the other hand, we need to focus more on the design of interpersonal interaction rather than structuring the program in critical thinking learning that pursues improving critical thinking ability. The result shows that cognitive achievement in interactive instruction is significantly higher than the achievement in structured instruction when the learning contents are to pursue critical thinking ability. But interpersonal interaction does not need to be the teacher-student interaction all the time. Student-student interaction was rather effective enough in this research. This gives us an important implication. One of the reasons that the teacher’s interaction with students may not be implemented in the best way in reality is because the teachers feel overloading for interaction with the students (Lao, 2002). Since the interaction is labor-intensive and therefore expensive, distance universities may not be able to finance this type of instruction for all students. If it is come to be understood that the effect of student-student interaction is not inferior to the effect of student-teacher interaction, it could save a substantial budget for hiring instructors or tutors. In this regard, findings in this research implicate that interactions among students need to be designed elaborately to relieve teacher’s overloading and to maximize learning effects at the same time.

Meanwhile, instructional method factor seems to have greater influence on student achievement than learner characteristics factor does. Perceived learner characteristics were analyzed a week prior to each program’s implementation in this research as well as in the pilot implementation. And the results showed no difference regardless of learner’s characteristics. It implies that a student’s learning pattern would be more influenced by instructional methods or teacher factors rather than by learner characteristics. This means that between variable (instructional method, different in group between) is more critically influential than within variable (learners’ characteristics, different within group) on student learning. Learning is, of course, influenced by the learner’s cognitive, emotional, and social characteristics. But understanding teaching and learning situation needs to be approached comprehensively. In any case, it’s impossible to instruct a perfectly homogenous group as there are students’ differences all the time. In spite of the existence of individual’s differences, some classes’ average GPA does go up, and popular instructors and successful programs exist. This indicates that there are certain effective teaching methods in a context. This makes us aware of the importance of instructional design (ID) for effective teaching method.

In conclusion, this research provides us a significant implication for understanding theoretical framework in distance education. As reviewed in Theoretical Background, the contrary and complementary relationship between structure and interaction has been pervasive in academia. The idea of when structure increases, interaction decrease, and vice versa, implicates one can be a substitute for another. This research shows empirical evidence that structure can be a substitute in a certain case although it’s not always the case. With regard to pedagogic optimum between two extremes, with an extremely structured course at one end and extremely interactive course at the other end, this paper suggests that the optimum should be located in structure side in receptive learning, whereas it should be moved a lot.
toward interaction side in critical thinking learning. The practical implication of this is that we need to consider the characteristics of learning contents first, such as receptive learning or critical thinking learning, in web-based distance learning program development, because the learning outputs are influenced more by characteristics of the learning contents, especially in deciding the priority between self-study mode with highly structured program and interactive mode with highly interpersonal interactive program, or how to combine each mode.

Finally recommendations for further research are suggested as follows: first, this research does not consider each structuring elements’ effect or influence on learning, although many structuring elements from previous literature were applied to structured instructional development. But there could be many different design types among structured instructional program. Analysis and comparison of each element’s influence on learning would be a good theme for further research. Second, more objective analysis of learner’s characteristics with verified measuring tools, including special characteristics such as learner autonomous, aptitude, or self-regulating ability, would give us more articulated implications. Third, further research might study on structuring of interaction. Stein et al.(2005) mentioned that the combination of learner-initiated interaction with instructor-initiated interaction built into the course explained more about satisfaction with perceived knowledge gained than satisfaction with structure alone. Thus, research on structuring of interaction will enhance our flexible understanding of two extreme concepts in distance education. Fourth, although basic principles of pedagogy is not easily changed, further study can explore if there is any difference in the results according to IT technology evolution which might have influence on the way of contents development and interaction than before. Finally, further research providing quantitative data from more cases with various learning contents could verify and generalize the findings of this research. Also research considering emotional or social evaluation as well as cognitive evaluation on various learning contents could be recommended for more comprehensive understanding in web-based distance education.

References


Preservice Teachers' Perceptions of Applets for Solving Mathematical Problems: Need, Difficulties and Functions

Wajeeh Daher
Department of Mathematics, Al-Qasemi Academy, Baka, Israel // wdaher@macam.ac.il

ABSTRACT
This research describes preservice teachers' perceptions of the use of applets in solving mathematical problems. It examines such perceptions regarding the role and functions that applets have in the problem solving context, the difficulties they encounter when using applets in this context and the necessity or need of their use. I attempt to categorize the perceived need in terms of the types of solution that were used by the participants and the difficulties that they met during solving mathematical problems. To analyze the participants' solutions, their difficulties and need for applets I use the content analysis method which enables to determine the presence of certain words or concepts within the text and the relation between these words or concepts. To discuss the participants' perception of the need for applets I used the activity theory frame which enables to analyze the participants' opinions by looking at their operations, actions and activity, and, at the same time, looking at their goals and motives. Though most of the participants thought that mathematical problems could be solved without applets, they emphasized the role of applets as fostering, facilitating and clarifying mathematical problems' statement and solution. The participants pointed at applets as tools which learners enjoy working with, so they will be encouraged to solve mathematical problems using them. What influenced the participants' perception of the need to use applets for solving mathematical problems were their ability to solve problems using them, their ability to solve problems without them and the type of difficulties faced by them during the solving process.

Keywords
Applets, Mathematical problems, Preservice teachers, Perceptions, Activity theory

Introduction

Background

The use of the internet as an educational medium is now rapidly expanding (Liaw, 2004). Sinclair (2000) points that the internet is becoming an intriguing environment for mathematics education by combining a variety of modes that can be applied to learning. Engelbrecht and Harding (2005) say that there are an increasing number of mathematical sites that use applets to enhance their pages with animated figures and interactive illustrations. They point that these sites include good mathematical applets that are visually of great value and that can be fruitfully used as educational tools. Applets are programs that require a WWW browser or other application to run. Usually these tools represent scientific concepts algebraically and visually, so they enable learners to perceive different representations of the same concept. Examples of mathematical applets can be found at many educational sites, for example the National Council of Teachers of Mathematics' (NCTM) site or the Center of Educational Technology' (CET) site, where the first site includes applets that treat different mathematical topics, while the second site includes applets that treat two topics: linear and quadratic functions. Usually an applet treats a specific topic, for example number factors, decimal fractions, the circle area, etc. So to cover a whole subject one needs different applets, for example one needs to work with applets that treat the following topics to cover the linear function subject: input-output, rate of change, transformations, transformations of straight lines which pass through the origin, addition of linear functions, subtraction of linear functions, forms of linear functions, and functions defined on intervals. The CET's site includes an applet for each topic mentioned above, where the topics are presented in a problem solving context. I've been introducing this environment for preservice teachers for almost 8 years now. Some of them find difficulties working with applets to solve mathematical problems, while others don't find such difficulties, some don't consider applets to be needed in solving mathematical problems, while others are attracted to these new tools and find them helpful in solving the problems. I have always asked myself about the factors that make some learners find difficulties while working with applets while others don't do so, and about the factors that make some learners approve working with applets while others think applets are not needed. This research came to answer these questions.

Literature review

Bolyard and Moyer (2003) say that students in middle grades sometimes need extra help and scaffolding to make the transition from concrete tasks to abstract concepts. They suggest to use contextualized problems with multiple...
representations that applets can provide as means to help perform this transition. Young (2006) provided a summary of the current literature on applets in mathematics education and described their benefits as found in various articles: (1) their availability on the internet and thus their free and ease of access (2) their focus on specific concepts (3) applets enable learners to do things that are not possible or easy with physical manipulatives, or pencil and paper (4) applets provide students with instantaneous and corrective feedback, so they fit inquiry-based learning and problem solving (5) applets provide multiple representations of a single concept at the same time; thus they can promote the transfer of knowledge from specific ideas to general knowledge (6) applets may be helpful for students with disabilities (7) applets increase motivation and attention in students as well as teachers.

The ability of applets to represent mathematical phenomena in multiple representations and to motivate students' learning of mathematics make them an appropriate tool for solving mathematical problems, especially word problems which students don't look forward to solve them. It may be expected that the ability of applets to offer multiple representations can introduce different types of solution for mathematical problems, for example: dynamic solution (applet only solution), blended solution (applet and algebraic). These main aspects of applets encourage exploring their use for solving mathematical problems which are usually solved in algebraic methods.

Eason and Heath (2004) said that applets enhance mathematics learning and added that they helped immensely when first introducing solution methods for discrete dynamical systems. They pointed that students using applets dealing with discrete dynamical systems could manipulate the parameters of a given problem and thus obtained a system appropriate for it. Then they could plot and visualize the behavior of the system. Eason and Heath described how the learning of the students was enhanced by working with an applet's window which computed the numerical solution of the system and with an explanation window. All these options enabled students to explore deeper concepts behind the behavior of the discrete dynamical systems.

Gadanidis, Gadanidis, and Schindler (2003) found that the use of applets was one of several factors; typically not the major factor, affecting the pedagogical thinking of preservice teachers when planning mathematics lessons that use applets. They mentioned the following factors as affecting the preservice teachers' pedagogical thinking: the mathematics' topic, their pedagogical beliefs, personal mathematics' knowledge, comfort with technology, classroom management concerns, past mathematics learning and teaching experiences, and recent teacher's educational experiences.

Smetana (2003) described how two 8th grade students and one 7th grade students worked with applets to solve balance problems. First they solved balance problems in worksheets, and then in applets. The students found working with applets difficult because they had hard time using the program. They also preferred working with the problems in the worksheet over the problems in the applets. Smetana commented on her experience saying that the kind of applets that she used would probably benefit younger students who hadn't been introduced to balance problems yet.

Gadanidis (2002) reports case studies of three 5th and 6th grades teachers and one 10th grade teacher who used applets in their teaching of mathematics. The 5th grade teacher worked with students on maximum area problems, first without applets and then with applets, and was surprised how all students, not only the top ones, could solve maximum area problems without much difficulty. The 6th grade teachers worked with their students on the same problems, using applets from the beginning. They reported that many of their students didn't know what to expect, thought that working with applets was for fun and didn't know how to solve the maximum area problems using them. The 6th grade teachers concluded that they would better first familiarize the students with the problems and then ask them to solve them with applets. The 5th and 6th teachers said that working with applets made them aware of the relation between perimeter and area, which they hadn't been aware of before. The 10th grade teacher reported that the class who worked with applets gained significantly higher scores than the class who didn't. The teacher also pointed that working with applets has shifted his attention from teaching individual and isolated effects of coefficients of quadratic equations to holistic and dynamic exploration of relationships between quadratic equations and graphs.

Theoretical framework

Activity theory

Activity theory is a cultural-historical theory of activity which was initiated by a group of Russian psychologists in the 1920s and 1930s. This approach was led by Lev Vygotsky (1896-1934) and his colleagues A. N. Leont'ev and A. R. Luria (Engestrom, 2003).
The basic structure of an activity is represented in Figure 1.

In this structure, an activity is composed of a subject and an object, mediated by a tool, where the subject is a person or a group engaged in an activity, while an object is held by the subject and motivates the activity, giving it a specific direction. The mediation can occur through the use of different types of tools, material as well as mental, including culture, ways of thinking and language (Mappin et al. 1999). According to Leont’ev (1979), analysis, using activity theory, takes into account three levels: analyzing the activity and its motive, analyzing the action and its goal, and analyzing the operation and its conditions. Leont’ev describes the actions as the activity components that translate it into reality. Operations refer to how actions can be done under objective circumstances. The structure that Leont'ev suggests for an activity is represented in Figure 2.

This structure will be used to explain the relations between the history of the participants, regarding their use of applets in learning mathematics, and their perceptions regarding the need for applets when solving mathematical problems.

**Areas of ICT integration in learning**

Wang and Woo (2007) say that, depending on the scope of content covered, ICT integration can happen in three areas: curriculum, topic and lesson. They further elaborate explaining that ICT integration into the curriculum requires ICT to support a more substantial amount of subject content which contains a number of topics in a specific discipline. ICT integration in a topic can be used to cover certain topics within a course. At the micro level, ICT is used to help explain specific knowledge units which can be covered within a single lesson. This structure of ICT integration will be used to categorize the functions perceived by the participants to be played by applets in mathematical problem solving.
Martínez-Santaolalla et al. (2005) describe three phases of learning that students should progress through in order to acquire true understanding of mathematics: action (involving physical and mental manipulations), abstraction, and reflection (when students consciously analyze their thinking). Martínez-Santaolalla et al. point that cycling through these phases time after time enables students to construct increasingly sophisticated mental models of the mathematical phenomena studied.

The participants in this research went thought all the previously mentioned three phases of learning when they solved problems from the CET site.

Rationale for the study

Preservice teachers should be exposed to technological tools that would help them in their future teaching. They should first be exposed to these tools as learners, so they would appreciate them as tools that facilitate learning, and thus, when they come to teach, they will introduce them to their students as objects with which they can learn. Technological tools can have various functions in the mathematics classroom; one of which is to solve mathematical problems with them, which is an important aspect of the students' mathematical activity (NCTM, 2000). In this research, I want to examine the work of preservice teachers with applets to solve mathematical problems and how they perceive this work. This perception can point to whether and how they will use the applets as teaching tools. This research will also help us better know how to introduce applets as solving tools to preservice teachers.

Designers would benefit from this research because they will become more knowledgeable regarding the difficulties that users experience when working with applets. This would help them design more efficient applets that students have less difficulties when working with them.

Young (2006) says that very little formal research has been performed on the effectiveness of applets. This research is interested to explore the perceptions of preservice teachers regarding the functions, effectiveness, and benefits of applets, while solving mathematical problems. Doing so, this research attempts to permeate a little researched field.

The research questions

1. What are the roles and functions of applets that preservice teachers perceive when using them to solve mathematical problems?
2. What are the difficulties faced by preservice teachers when solving mathematical problems with applets?
3. What is the relation between the perceived need for applets to solve mathematical problems, and the types of solutions and difficulties faced during the solving process?

The research method

The research setting and participants

The research was conducted in two mathematics didactics classes of preservice teachers who learned in an academic college of education. The preservice teachers were in their second year of study to major in mathematics and computer education. These classes were held in two consecutive academic years 2006-2007 and 2007-2008. The first class included 27 preservice teachers, while the second included 28 preservice teachers. The mean age of the participants was 20.7 years with standard deviation of 1.3. The didactics course introduces the preservice teachers to various technological methods of teaching mathematics: applets, video, spreadsheets and wiki. The didactics course follows a geometry course in which the preservice teachers are introduced to and work with applets.

The assignments

The participants solved two assignments with applets. The first assignment was a task in the geometry class, where the participants were required to solve a geometry problem using an applet. The second assignment was a task in the didactics course, where the participants were required to solve an Algebra problem about quadratic functions using
different applets which the CET’s site provided. The participants had two weeks to solve each assignment. They worked on the assignments in pairs or individually. The first assignment was given to the participants in the fifth and sixth weeks of the academic year, while the second assignment was given to the participants in the sixteenth and seventeenth weeks of the academic year.

The assignment that this research reports is the second assignment. In this assignment the participants were required to do the following: (1) solve an Algebra problem using the applets that accompany it (2) describe in detail their solving process, and (3) reflect on their work with the applet, emphasizing the difficulties that they met and their opinion regarding the need for applets in solving the mathematical problem they had solved.

The electronic book

The Algebra problems were part of an electronic book: Visual Math: Functions, at the CET site. This electronic book includes mathematical problems which can be solved using three types of applets.

1. An applet which is part of the problem, as figure 3 shows.

![Figure 3: an applet which is part of the mathematical problem](image)

Usually this type of applets enables the user to drag the red points in the figure to change its dimensions; in our case its length, width, diagonal and perimeter.
2. An applet which is designated to be used with all assignments in a specific topic that the mathematical problem is part of. For example, the applet “The difference equation” is an applet designated to be used in solving all the assignments in the mathematical topic “solving equations”. The applet's link lies at the upper left of the web page where the mathematical problem lies. Figure 4 shows this applet.

![Figure 4: An applet for all the assignments of the topic: ‘solving equations’](image)

This applet enables the learner to write the algebraic rules of any two functions and to get their difference function algebraically, visually and dynamically.

3. An applet which designated to be used with any assignment in the whole electronic book. This type of applet is in a box called ‘Tools’. This box can be accessed by clicking a button called ‘tools’ in the upper right of any web page of the electronic book. Figure 5 shows this box.

![Figure 5: the box of tools designated to be used throughout the whole electronic box](image)

The research tool for analyzing the data: content analysis

I used content analysis methodology to arrive at categories of applets' functions when used to solve mathematical problems, and categories of the difficulties faced by the participants in the solving process. Content analysis determines the presence of certain words or concepts within a text in order to analyze the meanings and relationships...
of such words and concepts. These meanings and relationships have usually inferences about the messages within the texts regarding the answers to the research questions.

Two experienced coders (the author is one of them) coded the participants’ solutions and reflections on their solutions, searching for categories of the participants’ perceptions regarding the role and functions of applets and of their difficulties while solving interactive mathematical problems. The agreement between the coders (Cohen’s Kappa coefficient) was 0.78 for functions of applets and 0.84 for faced problems). This type of analysis is familiar in examining perceptions of students (Wishart, 2000; Keller and Cernerud, 2002; Lizzio and Wilson, 2008)

Findings

Role and functions of applets when solving mathematical problems

The preservice teachers pointed at the following roles and functions which applets might fulfill in the mathematical problems’ context:
1. Representing the problem:
   - They help to design the story or the phenomenon in the mathematical problem.
   **Preservice teacher:** We started to build the playing fields that fulfill the conditions given in the problem.
   - They help students feel and see mathematical terms, objects and concepts.
   **Preservice teacher:** I can give this applet to middle school students to let them feel and see what areas and circumferences mean.
2. Enabling students to experiment and repeat their solution steps:
   - Applets help students repeat what they do till they see a pattern.
   **Preservice teacher:** I did some experiments and noticed the changes that the graphs underwent as a result of changing the functions' parameters. I continued my experiments till I arrived at mathematical relations.
3. Clarifying the story or phenomenon that the mathematical problem describes:
   - They help clarify mathematical conditions which the problem doesn't mention.
   **Preservice teacher:** When I read the problem I wondered what figure the hall should have, because that wasn’t mentioned in the text, but when I worked with the applet I knew that the hall should be a rectangle.
4. Influencing the nature of the mathematical problem's solution:
   - Applets enable students to solve mathematical problems graphically or using tables of values.
   **Preservice teacher:** Applets let us solve mathematical equations graphically by looking at the intersection points of two functions, or finding all the x values which have the same y in the table of values.
   - Applets enable us to solve mathematical problems in different methods.
   **Preservice teacher:** Applets provide us with another solving method, other than the algebraic solution.
5. Applets help to evaluate one’s solution:
   - They enable the learner to check the correctness of her solution of a mathematical problem, because she can compare her algebraic solution to the one obtained from the applet.
   **Preservice teacher:** We decided to solve the problem in two ways: algebraically and with the applet. Doing so, we could compare between the results obtained by the two methods. This enabled us to verify the correctness of our solutions.
6. Influencing the easiness of solving mathematical problems:
   - Applets help the learner solve mathematical problems easily and in a short time.
   **Preservice teacher:** I didn’t expect the applet to lessen the effort of solving the problem. The solution consisted of considering three cases for every part of the problem, so I had to solve algebraically a quadratic equation for every case. I dispersed with this solving process by working with the applet. I just wrote the functions and the applet gave me the intersection points, then I continued algebraically.
7. Influencing the motivation of solving mathematical problems:
   - Applets motivate students to solve mathematical problems.
   **Preservice teacher:** I liked working with the applet, and this motivated me to spend a long time solving the problem.
   - Applets make solving mathematical problems enjoyable.
   **Preservice teacher:** I’m sure that pupils in the middle school will enjoy solving problems with applets. I learned that from my own experience.
Table 1 describes the percentage of preservice teachers who mentioned each role or function.

<table>
<thead>
<tr>
<th>Role or function of the applet</th>
<th>Percentage of preservice teachers who mentioned the function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representing the problem</td>
<td>45.45%</td>
</tr>
<tr>
<td>Enabling students to experiment and repeat their solving steps</td>
<td>10.91%</td>
</tr>
<tr>
<td>Clarifying the story or the phenomenon that the mathematical problem describes</td>
<td>7.27%</td>
</tr>
<tr>
<td>Influencing the nature of the mathematical problem's solution</td>
<td>18.18%</td>
</tr>
<tr>
<td>Helping to evaluate one’s solution</td>
<td>18.18%</td>
</tr>
<tr>
<td>Influencing the easiness of solving mathematical problems</td>
<td>21.82%</td>
</tr>
<tr>
<td>Influencing the motivation to solve mathematical problems</td>
<td>10.91%</td>
</tr>
</tbody>
</table>

We see from table 1 that the applets' function which mostly drew the participants’ attention was the ability to represent mathematical problems in ways other than algebraic.

It should be noted that the participants pointed at applets as effective tools for teaching and learning mathematics. They did so while describing the various functions of applets, like representing problems, making the solution process easier, motivating students, etc.

**Difficulties encountered by the participants while solving mathematical problems**

The preservice teachers pointed at the following difficulties faced while solving mathematical problems with applets:

1. Operating problems:
   *Preservice teacher:* It was so difficult at the beginning for me to find the intersection points of two functions. I used to look at the table of values and change the steps between consecutive x values till arriving at the intersection points. Only at the end did I realize that it’s so simple to see the coordinates of the intersection points by clicking on them.

2. Mathematical problems:
   *Preservice teacher:* I had difficulty at the beginning remembering which formula I should use in order to compute the area of the triangle in the problem.

3. Matching an applet to a mathematical problem:
   *Preservice teacher:* After reading the problem, I worked with the applet that accompanies it, but I couldn’t solve the problem with that applet, so I went to ‘Tools’ to find what applet is appropriate for solving the problem. I found that the ‘binary’ applet could solve it, but it took me much effort and time.

4. Solving a problem with applets:
   *Preservice teacher:* It was so difficult at the beginning to solve the problem using the applet that accompanies it or the applet in the left column.

5. Matching the diagram which the applet describes with the story of the problem:
   *Preservice teacher:* The applet should have had two types of boxes; one with a cover and one without a cover, but it had only a box with a cover. This made it hard for me to continue my experimentation.

6. Getting precise mathematical values from an applet:
   - Getting exact solutions
     - (a) Difficulties faced because of the nature of mathematical objects:
       *Preservice teacher:* I had difficulty using the applet to arrive at the exact answer of the quadratic equation because the answer was an irrational number.
     - (b) Difficulties faced because of the nature of the applet's processes:
       *Preservice teacher:* I spent something like ten minutes to arrive at the correct answer because every tiny dragging changed the values of the area of the playing field.

7. Constructing or drawing shapes with required dimensions:
   - Drawing a shape with specified lengths:
     *Preservice teacher:* I had difficulty drawing halls that had the given lengths because the applet didn’t enable drawing
rectangles with big areas.
8. Seeing the mathematical values which the applet shows.

**Preservice teacher:** I couldn’t see the area of the playing field because the color of the area measurement was the same as the color of the playing field.

Table 2 describes the percentage of preservice teachers who mentioned facing each difficulty.

<table>
<thead>
<tr>
<th>Difficulty faced by the preservice teachers</th>
<th>Percentage of preservice teachers who mentioned facing the difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating difficulties</td>
<td>10.91%</td>
</tr>
<tr>
<td>Mathematical difficulties</td>
<td>45.45%</td>
</tr>
<tr>
<td>Matching an applet to a mathematical problem</td>
<td>18.18%</td>
</tr>
<tr>
<td>Solving a problem with applets</td>
<td>21.82%</td>
</tr>
<tr>
<td>The matching of the diagram which the applet describes with the story of the problem</td>
<td>7.27%</td>
</tr>
<tr>
<td>Getting precise mathematical values from an applet</td>
<td>10.91%</td>
</tr>
<tr>
<td>Constructing or drawing shapes with wanted dimensions</td>
<td>14.55%</td>
</tr>
<tr>
<td>Seeing the mathematical values which the applet shows</td>
<td>14.55%</td>
</tr>
</tbody>
</table>

It should be noted that only two participants didn’t report having difficulties while using applets to solve mathematical problems.

**Methods used by the preservice teachers to solve electronic mathematical problems**

The participants used three methods to solve mathematical problems: algebraic methods, solving with applets only, blending between algebraic methods and applets.

Solving with applets:
The participants usually started solving the mathematical problem by using the applet which accompanies the problem's text. When they didn’t succeed to solve it in this way, they tried applets designated for the whole topic that their problem was part of. When they didn’t succeed too, they tried one of the applets in the Tools box.

Algebraic solution:
The participants solved a full algebraic solution in two cases: when they could not solve a problem with any available applet, or when they wanted to find out if the result obtained with applets matches the result that can be obtained algebraically.

Blended solution:
Some participants could solve with applets just part of the problem but not all of it, so they solved the rest of it algebraically. Other participants used algebra and applets together all along the solving process.

**The need for applets to solve mathematical problems**

There were four main perceptions of the need to use applets in solving mathematical problems:
1. Mathematical problems can be solved without applets, but working with them would benefit pupils in the middle school because they would visualize the mathematical problem, and thus understand it. Those who had this perception could solve the mathematical problem easily in an algebraic method but said they enjoyed working with applets. Some of them found difficulty using the applet to construct shapes with exact lengths, but nevertheless enjoyed solving the problem with applets.
2. Mathematical problems can be solved without applets, but working with them would make the solution process easier, faster and more enjoyable. Those who had this perception could solve the mathematical problem algebraically. All of them also solved the
problem in a method which blended algebra with using an applet, or which used only an applet. Some of these participants had mathematical difficulties.

3. Mathematical problems can be solved without applets.

Those who had this perception could easily solve the mathematical problem algebraically. At the same time, they couldn’t solve all parts of the problem using one of the available applets.

4. Some problems can’t be solved without applets.

Those who had this perception couldn’t solve their problems without applets. At the same time, they had difficulty understanding the mathematical problem before working with the applet.

Discussion

Niess (2006) points that to be prepared to teach mathematics, teachers need an in-depth integrated knowledge of mathematics, teaching and learning, and technology. Working with applets to solve mathematical problems, the preservice teachers built such integrated knowledge, perceiving, on the way, different roles of applets and facing different difficulties. These two factors influenced their perceptions of the need for applets to solve mathematical problems. I analyze each of these perceptions starting from the perceived role of applets.

Role and functions of applets when solving mathematical problems

To categorize the functions of applets perceived by the participants, it's possible to go after the participants who talked about applets as tools for teaching and applets as tools for learning. Using the ICT integration model of Wang and Woo (2007), regarding the three areas of ICT integration in learning, we can categorize this perception of the role of applets as accepting the applets as tools to use in the whole curriculum, not just in a specific topic. Applets were further perceived by the participants as tools for evaluating their results. Talking about applets as tools for learning, in our case solving mathematical problems, the preservice teachers emphasized the representational aspect of applets. This emphasis can be explained by the influence of different representations on the understanding of different aspects of the problem, and thus the ability to approach its solution from a better stand. This finding about the emphasis on representations is supported by the literature, for example, Santos-Trigo and Cristóbal-Escalante (2008) pointed that the use of the Cabri-Geometry software was an important tool for students to construct dynamic representations of the problems. This happened because it was used to identify and examine different mathematical relations. On the other hand, Hwang et al (2007) found that elaboration in creativity was significantly correlated with multiple representation skills. Further, the preservice teachers pointed that working with applets makes the solution process more enjoyable, so the students can be motivated to solve more mathematical problems. The enjoyment and motivation aspects could be explained by the animation aspect of applets and by the interactivity between students and applets (Crisp, 2002).

We can say that the participants perceived applets as tools that could change the solution nature of mathematical problems, and, on the other hand, as tools that enrich the problem solving context.

Difficulties faced by the participants while solving mathematical problems with applets

We can categorize the difficulties faced by the participants while solving mathematical problems using applets into: (1) Operating difficulties (2) mathematical difficulties (3) matching difficulties, and (4) compatibility difficulties. Few participants reported having operating difficulties; besides, these difficulties weren’t difficulties of inability to operate an interface element, but inability to see a ‘hidden’ element, for example, to see that by clicking on the intersection point of two graphs, one can find the coordinates of that point.

Mathematical difficulties faced by the participants were of different types: (1) understanding the story of the mathematical problem (2) understanding the conceptions of the mathematical letters that inhibit the applet (3) finding a mathematical formula which could be used to solve the problem. These difficulties emphasize the advantage of applets in the middle school, because middle school students are sure to have more mathematical difficulties than the participants, thus they can be assisted by applets to overcome their mathematical difficulties of the first and third types.
Matching difficulties were of three types: (1) finding an applet appropriate for solving a given problem (2) finding the relation between a diagram in the applet and the verbal text of the problem and (3) knowing how to solve the mathematical problem using an applet.

The first matching problem faces the student before working with an applet to solve the mathematical problem, the second matching problem faces the student at the beginning of the solving process, and the third matching problem faces the student throughout the solving process.

The matching difficulties can be explained by the short history of the participants regarding working with applets to learn mathematics, and thus they will be overcome as the learner becomes experienced in working with mathematical applets.

Compatibility difficulties were difficulties that involved the inability to get specified mathematical values or precise mathematical values using an applet, or to construct with the applet geometric shapes that have specified lengths. Generally, compatibility problems can’t be overcome by students, so they should be a challenge for software designers.

The need to use applets for solving mathematical problems

Though the preservice teachers perceived various functions and benefits of applets, as tools for solving mathematical problems, most of them thought that they can solve the problems without them. Thinking so they emphasized that, though solving mathematical problems wouldn't necessarily need applets, working with them would make the solving process easier, faster, and more enjoyable; especially for middle school students. They also pointed at the help they can get from working with applets to solve mathematical problems: "applets help us understand the problem, construct it, and have ideas how to solve it". These perceptions of the participants agree with the findings of Martinez-Santaolalla et al. (2005) who found that dynamic geometry software facilitates and enhances the learning of students.

We can categorize the perception of the participants regarding the need for applets as tools for solving mathematical problems into three categories: (1) applets aren’t necessary for solving mathematical problems (2) applets are necessary for solving mathematical problems and (3) applets aren’t necessary for solving mathematical problems, but they can enrich the solving process.

Those that considered applets unnecessary to solve mathematical problems had two history events: (1) They could solve easily the mathematical problem algebraically, but (2) they couldn’t solve some parts of the problem using only applets or in a blended method.

Those that considered applets necessary for solving mathematical problems or some mathematical problems had the following couple of history events: (1) Applets helped them throughout the entire process of solving the mathematical problem, and (2) They enjoyed working with applets to solve mathematical problems.

Those that considered applets to be unnecessary for solving mathematical problems, but considered using them as enriching the solution process, had the following couple of history events: (1) They could solve the mathematical problem working only with applets, and (2) They could easily solve the mathematical problem algebraically. Some of them also reported enjoying working with applets.

These relations between the history events and the perceptions of the preservice teachers, regarding the need for applets when solving mathematical problems, can be explained using the frame of activity theory as formulated by Leont'ev (1979).

Table 3 shows the activity of a preservice teacher, when solving a mathematical problem with applets, using Leont'ev's framework.

Those that could perform the actions and operations of the activity successfully, understanding, manipulating and utilizing the graphical signs, and enjoying this action and these operations, perceived the applets to be needed when solving mathematical problems, while those that couldn’t perform the actions and operations of the activity
successfully perceived the applets to be unnecessary for solving mathematical problems. The third party who succeeded to perform the action and operations of the activity, but succeeded easily to perform other operations that didn't follow the conditions of the activity (like working with algebraic only signs), perceived the applets as not necessarily needed but favored when solving mathematical problems.

Table 3: describing the work of the participants with applets using the activity theory frame

<table>
<thead>
<tr>
<th>Activity and Motive</th>
<th>Action and goal</th>
<th>Operations and conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studying mathematics in an electronic book, having the motive to learn mathematics with technological tools that enable multi-representations.</td>
<td>Attempting to solve mathematical problems using applets or in a blended method: applets with algebra.</td>
<td>Working with graphic signs or both graphic and algebraic signs using applets</td>
</tr>
</tbody>
</table>

Tools are a very important component in activity theory. They mediate between the subject and the object of the activity. In our case, the subject of the activity is the student, while the object is solving mathematical problems with technological means, and the tool is the applets. We can say that the participants, perceiving the applet as a solving tool, differentiated between the following types of tools: (1) a tool that can solve problems which can’t be solved otherwise (2) a tool that can solve all mathematical problems (3) a tool that enriches the solving process but can’t solve all mathematical problems, and (4) a tool which is difficult to use to solve mathematical problems. Perceiving applets as tools from the first and the second types, the preservice teachers considered applets to be needed for solving mathematical problems. Perceiving applets as tools from the fourth type, the preservice teachers considered applets to be unneeded for solving mathematical problems, while perceiving applet as tools from the third type, the preservice teachers considered applets to be unneeded to solve mathematical problems, but working with them would help the learner during one or more stages of the solving process.

We can say that actions performed with applets, with a specific goal in mind, and their related operations, with specific conditions like the applet design, influenced the preservice teachers’ perception of applets as a cognitive tool which is needed when solving mathematical problems.

Some researchers described the importance of design issues of educational software (for example Yerushalmy, 1999; Misanchuk and Hunt, 2005). This research’s findings point at design deficiencies which can be avoided. For example, some participants claimed that some applets don’t enable the construction of geometric shapes with big values, or that the different colors of the applet's interface don’t allow seeing some measurements of the shapes in the applet. So the finding of this research can be beneficial for software designers too.

Conclusions

The participants perceived different roles and functions of applets when solving mathematical problems. They enjoyed solving problems using applets despite the fact that they faced different difficulties while doing so.

Most of the participants thought that mathematical problems could be solved without applets, but emphasized the role of applets as fostering, facilitating and clarifying mathematical problems' statement and solving. At the same time they pointed at applets as tools which learners enjoy working with, so they will be encouraged to solve mathematical problems using them. This means that the preservice teachers will probably use applets in their future teaching. They will do so when they want to foster and facilitate their students' learning or to make them enjoy the mathematical problem solving.

Factors that influenced the participants’ perception of the need to use applets for solving mathematical problems are their ability to perform successfully the activity and its related actions and operations, and the type of difficulties which hindered this performance. This says that preservice teachers should first be introduced to applets that don't have operating or compatibility problems, so they would enjoy their work with applets and not confront difficulties that lessen their confidence in applets as solving tools. This would make the preservice teachers look at applets as worth-to-work-with tools.

Designers of educational applets can benefit from this research, taking into consideration its findings, to make efficient design decisions.
References


Street Gang: The Complete History of Sesame Street
(Book Review)

Reviewer:
Dale Erlandson
Graduate Student
Department of Journalism and Mass Communication
University of Wisconsin-Madison
erlandsondn@gmail.com

Textbook Details:
Street Gang: The Complete History of Sesame Street
Michael Davis
ISBN 978-0670019960
18 chapters, 384 pages

For those people raised on Ernie, Bert and Oscar the Grouch, a witty, nostalgic return to the early days of Sesame Street is exactly what the 21st century has been missing. In Street Gang, Michael Davis attempts to weave together the professional and personal histories that launched a new era in children’s television. While the bulk of Davis’ narrative is framed by Jim Henson’s all-too-short life, much of the action revolves around Joan Cooney, the driving force behind Children’s Television Workshop. After a brief look at the Henson funeral, Davis retreats to the 1930s to trace the childhood dreams that guided a handful of performers, educators and activists to the 1970 debut of Sesame Street. After a shaky (and bewilderingly Captain Kangaroo-focused) start, Street Gang reaches firmer ground when all the personalities, including the Muppets, come together through Cooney and the organizations behind Sesame Street – namely, the Ford Foundation, the Carnegie Corporation, and the United States Office of Education. Only after the book begins to refer back to information introduced in the first three chapters, can the reader begin to make sense of the multitude of personalities and projects Davis’ whirlwind stirs up.

The book is the result of copious personal interviews with Sesame Street’s cast, creators and production crew, supplemented with newspaper and magazine commentary on the talent behind 1970s children’s television. Unfortunately for the historian, the interviews are poorly cited and the footnotes are sparse. The citations are subject to minor faults, such as missing ellipses when sentences are dropped, but the number of sources referenced is impressive and Davis has clearly tried to amass enough viewpoints to get the story right.

It is a pity, then, that Davis’ writing tends so much toward speculative-sounding detail. While minutia such as “streams of soft light on her pillow” (page 11) or the “feel [of] a solitary bead of perspiration run[ning] down his lower spine and trick[ling] to his waist, a sweat spot developing on his blue Oxford cloth shirt, just above the belt line” (page 49) might pass in a novel, they are unnecessary for readers who will no doubt bring their own nostalgia. That being the case, Davis’ unwavering forgiveness for the personality faults of those involved in Sesame Street seems gratuitous. So does his skimming over any unhappiness that may have arisen in production, at least until Henson’s death in 1990. In fact, it is hard to know for whom Davis is writing – while always engaging, Street Gang has less narrative than typical popular, light reading; but it is too poorly sourced for the serious historian of communications or education.

The biggest weakness of the Street Gang is its scope. Davis introduces a new personality – along with a full personal history and accompanying anecdotes – every few pages. The narrative is replete with temporal jumps that are nearly impossible to follow, and the attempt to cater to Sesame Street aficionados with trivia is botched by Davis’ insistence on inserting superfluous non sequiturs.
But for fans with the patience to simply ride along, it’s all there: how the name *Sesame Street* was picked over *123 Avenue B*; how the street-scene with its Brownstone and front steps were conceived; the revulsion some felt for the excessive juvenile simplicity of the now-beloved theme song.

At his best, Davis does an excellent job of contextualizing the show in a tumultuous time. The planning stages, spanning the 1950s and 1960s, saw most of the major talent brush with war in Korea or Vietnam. While *Sesame Street* itself never dealt with war, much of the show’s creative talent did find their professional lives interrupted and re-organized. For many who were originally hesitant about children’s educational television, the era’s social turbulence instilled a commitment to personal ideals and to the aims of *Children’s Television Workshop*. The start talent of the planning stages broke the misogynistic norms of corporate culture. In its first two decades, *Sesame Street* was both a reaction to and a target of the turbulence that defined the movements for civil, women’s, Hispanic and disabled rights. Davis deftly places the show’s development in the contexts of television and education history, which both weathered rapid technological and philosophical evolutions in the 1970s and 1980s.

While Davis makes clear the cronyism that brought many of the shows performers and producers together, he also reveals how the upper-middle-class progressive New Yorkers of the time did not have to look far for the issues they felt quality children’s television should address: writers worked in an actress’ pregnancy to address sibling rivalry and breast feeding; a producer’s son introduced both himself and Down syndrome to the neighborhood; and, all too soon, a spate of heart attacks, cancer and addiction forced *Sesame Street* residents, viewers and collaborators to cope with death.

The last third of *Street Gang* is a depressing downward spiral for *Children’s Television Workshop*, beginning with the death of Sesame Street general store owner Mr. Hooper and ending with the sale of the Muppets to the Walt Disney Company. For those who loved the early days of *Sesame Street*, Davis provides a refreshingly scathing review of the corporate mentality that took over after the horrific rise of Barney crippled *Sesame Street* ratings. Far less satisfying is Davis’ apparent satisfaction with the final result of the various revamp attempts that has left the show dominated by the gratingly high-pitched vapidities of Elmo and Zoe.

Davis’ work iterates that Henson, along with most of the other creative minds behind the 1960s re-think of children’s television, were dedicated to appealing to children and adults alike. The Muppets were to entertain parents as much (if not more) than they did children. As Davis suggests, *Avenue Q* – and not the overhauled *Sesame Street* – comes far closer to what Henson might have had in mind for the future of puppetry.