

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/220374450>

Designing Collaborative E-Learning Environments based upon Semantic Wiki: From Design Models to Application Scenarios.

Article in *Educational Technology & Society* · January 2011

Source: DBLP

CITATIONS

38

READS

4,102

3 authors, including:



Mingkai Dong

Siemens

8 PUBLICATIONS 121 CITATIONS

[SEE PROFILE](#)



Ronghuai Huang

Beijing Normal University

370 PUBLICATIONS 10,191 CITATIONS

[SEE PROFILE](#)

Journal of

*Educational Technology
& Society*

Published by

International Forum of Educational Technology & Society



**October 2011
Volume 14 Number 4**

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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to *Educational Technology & Society* and three months thereafter.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; **Nian-Shing Chen**, National Sun Yat-sen University, Taiwan.

Editors' Advisors

Ashok Patel, CAL Research & Software Engineering Centre, UK; **Reinhard Oppermann**, Fraunhofer Institut Angewandte Informationstechnik, Germany

Editorial Assistant

Barbara Adamski, Athabasca University, Canada.

Associate editors

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; **Kuo-En Chang**, National Taiwan Normal University, Taiwan; **Ni Chang**, Indiana University South Bend, USA; **Yam San Chee**, Nanyang Technological University, Singapore; **Sherry Chen**, Brunel University, UK; Bridget Cooper, University of Sunderland, UK; **Darina Dicheva**, Winston-Salem State University, USA; **Jon Dron**, Athabasca University, Canada; **Michael Eisenberg**, University of Colorado, Boulder, USA; **Robert Farrell**, IBM Research, USA; **Brian Garner**, Deakin University, Australia; **Tiong Goh**, Victoria University of Wellington, New Zealand; **Mark D. Gross**, Carnegie Mellon University, USA; **Roger Hartley**, Leeds University, UK; **J R Isaac**, National Institute of Information Technology, India; **Mohamed Jemni**, University of Tunis, Tunisia; **Mike Joy**, University of Warwick, United Kingdom; **Athanasis Karoulis**, Hellenic Open University, Greece; **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; **Jimmy Ho Man Lee**, The Chinese University of Hong Kong, Hong Kong; **Ruddy Lelouche**, Universite Laval, Canada; **Tzu-Chien Liu**, National Central University, Taiwan; **Rory McGreal**, Athabasca University, Canada; **David Merrill**, Brigham Young University - Hawaii, USA; **Marcelo Milrad**, Växjö University, Sweden; **Riichiro 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; **Jose A. Pino**, University of Chile, Chile; **Thomas C. Reeves**, The University of Georgia, USA; **Norbert M. Seel**, Albert-Ludwigs-University of Freiburg, Germany; **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; **Slavi Stoyanov**, Open University, The Netherlands; **Timothy Teo**, Nanyang Technological University, Singapore; **Chin-Chung Tsai**, National Taiwan University of Science and Technology, Taiwan; **Jie Chi Yang**, National Central University, Taiwan; **Stephen J.H. Yang**, National Central University, Taiwan.

Assistant Editors

Yuan-Hsuan (Karen) Lee, National Chiao Tung University, Taiwan.

Executive peer-reviewers

<http://www.ifets.info/>

Supporting Organizations

Centre for Research and Technology Hellas, Greece
Athabasca University, Canada

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

Educational Technology & Society is abstracted/indexed in Social Science Citation Index, Current Contents/Social & Behavioral Sciences, ISI Alerting Services, Social SciSearch, ACM Guide to Computing Literature, Australian DEST Register of Refereed Journals, Computing Reviews, DBLP, Educational Administration Abstracts, Educational Research Abstracts, Educational Technology Abstracts, Elsevier Bibliographic Databases, ERIC, Inspec, Technical Education & Training Abstracts, and VOCED.

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 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, please contact kinshuk@ieee.org (Subject: Submission for Educational Technology & Society journal).

Journal of Educational Technology & Society

Volume 14 Number 4 2011

Table of contents

Special issue articles

| | |
|---|-------|
| Guest Editorial – Advanced Learning Technologies <i>Ignacio Aedo, Mohamed Jemmi, J. Michael Spector and Larisa Zaiceva</i> | 1-1 |
| Design and Implementation of a 3D Multi-User Virtual World for Language Learning <i>María Blanca Ibáñez, José Jesús García, Sergio Galán, David Maroto, Diego Morillo and Carlos Delgado Kloos</i> | 2-10 |
| Language Technologies to Support Formative Feedback <i>Adriana J. Berlanga, Marco Kalz, Slavi Stoyanov, Peter van Rosmalen, Alisdair Smithies and Isobel Braidman</i> | 11-20 |
| Designing for Automatic Affect Inference in Learning Environments <i>Shazia Afzal and Peter Robinson</i> | 21-34 |
| A Learning Content Authoring Approach based on Semantic Technologies and Social Networking: an Empirical Study <i>Saša Nešić, Dragan Gašević, Mehdi Jazayeri and Monica Landoni</i> | 35-48 |
| Designing Collaborative E-Learning Environments based upon Semantic Wiki: From Design Models to Application Scenarios <i>Yanyan Li, Mingkai Dong and Ronghuai Huang</i> | 49-63 |
| A Workflow for Learning Objects Lifecycle and Reuse: Towards Evaluating Cost Effective Reuse <i>Demetrios G. Sampson and Panagiotis Zervas</i> | 64-76 |

Full length articles

| | |
|--|---------|
| Jordanian Pre-Service Teachers' and Technology Integration: A Human Resource Development Approach <i>Jamal Abu Al-Ruz and Samer Khasawneh</i> | 77-87 |
| An Individualized e-Reading System Developed Based on Multi-representations Approach <i>Chien-Chuan Ko, Chun-Han Chiang, Yun-Lung Lin and Ming-Chung Chen</i> | 88-98 |
| Guessing, Partial Knowledge, and Misconceptions in Multiple-Choice Tests <i>Paul Ngee Kiong Lau, Sie Hoe Lau, Kian Sam Hong and Hasbee Usop</i> | 99-110 |
| Computer-mediated Counter-Arguments and Individual Learning <i>Jack Shih-Chieh Hsu, Hsieh-Hong Huang and Lars P. Linden</i> | 111-123 |
| Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems <i>Yi-Hsuan Lee, Yi-Chuan Hsieh and Chia-Ning Hsu</i> | 124-137 |
| Using Wikis for Learning and Knowledge Building: Results of an Experimental Study <i>Joachim Kimmerle, Johannes Moskaliuk and Ulrike Cress</i> | 138-148 |
| A Constructivist Approach for Digital Learning: Malaysian Schools Case Study <i>Waleed H. Sultan, Peter Charles Woods and Ah-Choo Koo</i> | 149-163 |
| Ubiquitous English Learning System with Dynamic Personalized Guidance of Learning Portfolio <i>Ting-Ting Wu, Tien-Wen Sung, Yueh-Min Huang, Chu-Sing Yang and Jin-Tan Yang</i> | 164-180 |
| A New Approach Toward Digital Storytelling: An Activity Focused on Writing Self-efficacy in a Virtual Learning Environment <i>Yan Xu, Hyungsung Park and Youngkyun Baek</i> | 181-191 |
| Exploring Gender Differences in SMS-Based Mobile Library Search System Adoption <i>Tiong-Thye Goh</i> | 192-206 |

| | |
|--|---------|
| Time-Quality Tradeoff of Waiting Strategies for Tutors to Retrieve Relevant Teaching Methods <i>Wen-Chung Shih, Shian-Shyong Tseng, Che-Ching Yang and Tyne Liang</i> | 207-221 |
| Self-efficacy in Internet-based Learning Environments: A Literature Review <i>Chin-Chung Tsai, Shih-Chyueh Chuang, Jyh-Chong Liang and Meng-Jung Tsai</i> | 222-240 |
| Comparison of Web 2.0 Technology Acceptance Level based on Cultural Differences <i>Sun Joo Yoo and Wen-hao David Huang</i> | 241-252 |
| Factors Affecting Faculty Web Portal Usability <i>Rex P. Bringula and Roselle S. Basa</i> | 253-265 |
| Mining Learning Preferences in Web-based Instruction: Holists vs. Serialists <i>Natalie Clewley, Sherry Y. Chen and Xiaohui Liu</i> | 266-277 |

Book review(s)

| | |
|---|---------|
| Teaching Online: Tools and Techniques, Options and Opportunities (Nicky Hockly and Lindsay Clandfield) <i>Reviewer: David Leal Cobos</i> | 278-279 |
|---|---------|

Guest Editorial –Advanced Learning Technologies

Ignacio Aedo¹, Mohamed Jemmi², J. Michael Spector³ and Larisa Zaiceva⁴

¹DEI Lab, Universidad Carlos III de Madrid, Leganés, Spain // ²University of Tunis, Tunisia // ³University of Georgia, Athens, GA, USA // ⁴Riga Technical University, Latvia // ignacio.aedo@uc3m.es // Mohamed.jemmi@fst.rnu.tn // mspector@uga.edu // Larisa.Zaiceva@rtu.lv

ICALT, the International Conference on Advanced Learning Technologies, brings together researchers working on different disciplines related to the design, development, use and evaluation of technology-enhanced learning environments and to devising the new technologies that will be the foundation of the next generation of e-learning systems. In the 2009 and 2010 editions, celebrated in the wonderful cities of Riga (Latvia) and Sousse (Tunisia), three hundred researchers discussed about several topics related to the technology in the educational processes. ICALT 2009 received 310 submissions (266 full papers, 35 short papers and 9 posters) and 73 of them were accepted as full papers (27.44%). The acceptance rate of full papers of ICALT 2010 was 31.03% and 302 papers were submitted to the conference (258 full papers, 36 short papers and 8 posters). Editors of this special issue selected a number of full papers of both editions that got the highest scores during the conference review processes. These papers went through a peer review process for this special issue. 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, 6 papers were finally selected to illustrate some of the main advances presented in both editions of the conference.

The first paper of this special issue is “Design and Implementation of a 3D Multi-User Virtual World for Language Learning”, authored by Ibañez et al, shows an approach to foster communication skills within a 3D multi-user virtual world with minimum teacher’s help. In the second paper, “Language Technologies to Support Formative Feedback”, Berlanga et al propose the automatic development of conceptual maps of student works to ascertain learners’ progress and identify remedial actions. Next paper, “Designing for Automatic Affect Inference in Learning Environments” from Afzal and Robinson, sets a discussion of motivational and methodological issues involved in automatic affect inference in learning technologies. Semantic web technologies and web 2.0 are two of the trending topics that provide mechanisms to improve the learning processes and the two following papers apply them in different approaches: while the fourth paper, “A Learning Content Authoring Approach based on Semantic Technologies and Social Networking: an Empirical Study” from Nešić et al, uses these technologies to improve the authoring of learning contents, the fifth one, “Designing Collaborative E-Learning Environments based upon Semantic Wiki: From Design Models to Application Scenarios” from Li et al, makes use of them to facilitate collaborative knowledge construction and maximize resource sharing and utilization. Finally, Sampson and Zervas in their paper, “A Workflow for Learning Objects Lifecycle and Reuse: Towards Evaluating Cost Effective Reuse”, introduce a workflow for learning objects lifecycle that can support their reuse and a set of metrics for cost effective of their reuse.

This special issue tries to provide to the reader a broad panoramic of some working areas of ICALT. We hope you enjoy this special issue and that you explore more contributions to this research area in next ICALT conferences.

We recognize the contribution of reviewers: Nian-Shing Chen, Kinshuk, Demetrios G. Sampson, and Telmo Zarraonandia.

Design and Implementation of a 3D Multi-User Virtual World for Language Learning

María Blanca Ibáñez^{*}, José Jesús García, Sergio Galán¹, David Maroto, Diego Morillo and Carlos Delgado Kloos

Universidad Carlos III de Madrid, Dpto. Ingeniería Telemática, E-28911 Leganés, Madrid, Spain // ¹School of Arts and Communications (K3), Malmö University, 205 06 Malmö, Sweden // mbibanez@it.uc3m.es // rueda@it.uc3m.es // sergiogalan@eclectico.net // dmaroto@it.uc3m.es // dmorillo@it.uc3m.es // cdk@it.uc3m.es

^{*}Corresponding author

ABSTRACT

The best way to learn is by having a good teacher and the best language learning takes place when the learner is immersed in an environment where the language is natively spoken. 3D multi-user virtual worlds have been claimed to be useful for learning, and the field of exploiting them for education is becoming more and more active thanks to the availability of open source 3D multi-user virtual world development tools. The research question we wanted to respond to was whether we could deploy an engaging learning experience to foster communication skills within a 3D multi-user virtual world with minimum teacher's help. We base our instructional design on the combination of two constructivist learning strategies: situated learning and cooperative/collaborative learning. We extend the capabilities of the Open Wonderland development toolkit to provide natural text chatting with non-player characters, textual tagging of virtual objects, automatic reading of texts in learning sequences and the orchestration of learning activities to foster collaboration. Our preliminary evaluation of the experience deems it to be very promising.

Keywords

3D virtual learning environment, Learning system architecture, Technology-enhanced language learning, Open Wonderland

Introduction

One of the best ways to learn a foreign language is to be exposed to real situations in which it must be used to communicate (Genesee, 1985; Nieminen, 2006). Considerable advantages can be obtained by introducing collaborative activities (Zhang, 2010), promoting the participants' interaction with the environment and other members of the community. Nevertheless, the context must be somehow controlled; otherwise boredom or frustration might impede learning (Csikszentmihalyi, 1990).

A sound alternative to get the required level of linguistic immersion without losing control over the learning process are 3D multi-user virtual worlds (3DVWs). A 3D multi-user virtual world provides a shared, realistic, and immersive space where learners, by means of their avatars, can explore, interact, and modify the world (Bell, 2008; Calongne, 2008; Dalgarno & Lee, 2010; Dickey, 2005; Dillenbourg et al., 2002; Eschenbrenner et al., 2008; Girvan & Savage, 2010; Kalloni & Sampson, 2010). Furthermore, 3DVWs offer a rich environment in which learners can strongly interact with each other, increasing student's motivations for language learning (Andreas et al., 2010; Chittaro & Ranon, 2007; Hendaoui et al., 2008; Kluge & Riley, 2008; Lee, 2009).

Being immersed in a real environment and being able to interact with members of the educational community is not enough to learn a new language. As in any learning process, instructional design that focuses on specific learning outcomes is very important. In relation to learning outcomes, *The Instituto Cervantes in the Common European Framework of Reference for Languages* states that the "communicative language competences are put into operation with the completion of various activities that include language comprehension, expression and interaction. These activities can be classified as passive (reading, listening) or active (writing, speaking)" (Instituto Cervantes, 2002). When students communicate in a foreign language, they should demonstrate literacy in all those four essential skills (Hinkel, 2006; Nation & Newton 2009).

Our project is inspired by Language Learning projects that already use 3D Virtual Reality Technologies (Avatar Languages, 2009; Three Immersions, 2008; Koenraad, 2008; Shih & Yang, 2008) that simulate real environments and in some cases, real situations to promote speaking skills. We conceive the 3D learning system as a whole, as an integrated set of technological and pedagogical issues that are tightly related to one another, having to be dealt with

them independently but under a unifying light. This dual nature of our work will be reflected in this paper, in which we will describe both, the didactical developments as conceived, in the first place, and how they eventually have been brought into life by means of existing 3D technologies enhanced with our own developments. An analysis of existing related work will complete this schema. Finally we present a preliminary evaluation of our learning environment in terms of motivation, immersion, and participation in collaborative activities.

Related Work

Current instructional design models encourage active, rather than passive learning; they are based on constructivist theories whose central assumption is that humans create knowledge as opposed to acquiring it (Dewey, 1916; Ertmer & Newby, 1993; Vrasidas, C., 2000). Within the constructivist theory, there are two prominent schools: personal constructivism and social constructivism (Vrasidas, 2000). The former states that knowledge is constructed in the head of the learner, these principles follow Piaget theories. The latter assumes that knowledge is constructed in communities of practice, through social interaction (Vygotsky, 1978). Both schools emphasize the influence that the environment has on learners (Jonassen, 1994; Wilson, 1997).

Nowadays, information and communication technologies provide mechanisms to design and develop environments which facilitate the construction of knowledge and support personal and social constructivism (Perkins, 1992). Among the emerging technologies that can be used for distance education, 3DVWs is the one where it is possible to deploy truly immersive spaces fostering learner's imagination with possibilities of interaction with the environment, the objects and other community members through avatars (Bell, 2008; Calongne, 2008; Dillenbourg et al., 2002; Eschenbrenner et al., 2008; Girvan & Savage, 2010; Kalloni & Sampson, 2010). Thus, 3DVWs have offered, from their very beginning, an excellent place for learning and teaching and some authors have issued guidelines for using the constructivist approach on them (Chittaro & Ranon, 2007; Dillenbourg et al., 2002; Huang et al., 2010). The principles stated suggest the use of visual elements of 3DVWs to immerse students into a situation where the problem to be solved is presented in a natural way. Besides, 3D objects and non-player characters (NPCs) are used as instruments of transmitting information and as tools for building knowledge as it is required by constructivist principles. Projects following these guidelines include those on 3D simulation, public events organization and collaboration (Dalgarno & Lee, 2010; Livingstone & Kemp, 2008).

- *3D simulations.* This family of projects immerses participants into learning situations where they can practice in a safe environment with possibilities of having an individualized feedback. Among the projects that use full interactive simulations, it is worth mentioning the "Genome Island" (Clark, 2009), where visitors find interactive versions of classical genetic experiments, or "The Heart Murmur Sim" (Boulos, 2007), a training space where learners should diagnose the illness of patients. In the context of learning foreign languages, simulations involve recreation of real-life situations to promote students engagement (Shih & Yang, 2008) or the recreation of English-speaking town where students can have rich conversations not only with native speakers but also with their peers (Koenraad, 2008).
- *Organization of public events.* These applications use 3DVWs as meeting points or as mediums to explore learning environments. This kind of activities are usual in Second Life, with examples like the "New Media Consortium Campus" (Linden Labs, 2006), but not exclusive of that platform: the project MiRTLE ("A Mixed Reality Teaching and Learning Environment") (Callaghan et al., 2008), developed over Open Wonderland, is to be highlighted. Even if 3DVWs are one of the richest interfaces to be used for language learning, they are usually complemented with other tools and resources, like websites, audio chats, shared blackboards, etc. That is the case of the "3jSchool Chinese Language" (Three Immersions, 2008), whose virtual world, conceived for learning Mandarin, includes additional multimedia materials used in scheduled learning sessions.
- *Collaboration.* This last family of projects represents the essence of social constructivism, the possibility of creating knowledge within a learning community. Greenbush Edusim project (Greenbush Education Service Center, 2007) is an application in this category where students collaboratively build objects, as tangible knowledge.

A combination of these approaches is possible and even convenient for a foreign language learning environment.

In terms of learning strategies based on constructivism, instructional designers on 3D virtual learning environments have several possibilities (Girvan & Savage, 2010; Huang et al., 2010), here we survey some of them:

- *Situated learning.* It states that knowledge should be presented in an authentic context where it is possible social interaction (Dewey, 1916; Ertmer and Newby, 1993). 3DVWs enable deployment of simulations in realistic-looking environments.
- *Role playing.* Learners can assume different characteristics and personalities through their avatars (Holmes, 2007).
- *Cooperative/Collaborative learning.* 3DVWs can be seen as meeting points where learners can be aware of presence of peers, collaborate in building knowledge and communicate through the tools provided on these worlds (Chittaro & Ranon, 2007).
- *Problem-based learning.* 3DVWs allow to present ill-structured problems to solve, one of the principles of constructivism (Jonassen, 1994).
- *Creative learning.* H.M. Huang at al. (Huang et al., 2010) defends 3DVWs as environments that promote imagination and thus creativity.

If using 3D environments for teaching and learning seems a very sound option in general, when we focus on language learning the possibilities are really promising: 3DVWs become the ideal environment for deep linguistic immersion and realistic situated learning, without the need to travel to the places where the language to be learned is spoken. For the time being, our system includes 3D simulation of real conversations in downtown Madrid and group working specifically designed to enforce oral communication and information sharing. We base our instructional design on the combination of two constructivist learning strategies: situated learning and cooperative and collaborative learning.

All the learning foreign language projects reported concentrate their efforts in developing speaking skills but lacks of mechanisms to foster collaboration (Kreijns et al., 2003). In order to overcome these problems, the proposed system aims to develop the four communication skills orchestrated in a collaborative activity to achieve a final common goal.

Case study

The proposed learning experience takes place on a 3D multi-user virtual world that imitates cultural sights of Madrid in which a community of learners experience auditory and visual immersion. The scenery is filled with information about the life and work of D. Velázquez, one of the most important painters of Spain. Activities are designed to stimulate learners' imagination, to motivate them to acquire knowledge and to promote collaboration. Learners represented by customized avatars of their choice, should explore freely the environment looking for information to achieve a final goal: to get access to The Prado Museum.

In our 3D learning scenario, the activities are structured as the interaction of avatars with 3DVWs elements: synthetic environment, 3D objects, NPCs and other avatars. They are designed to develop and practice the skills involved in learning a foreign language.

- *Reading skills:* Reading skills are promoted through information tagged to 3D objects included in the scenario. When one of these objects is selected, its name appears along with practical information (reading comprehension). For instance, associated with the street names are written anecdotes about events that occurred there in the time of Velázquez.
- *Listening skills:* Listening skills are encouraged through interaction with 3D objects and NPCs. Some 3D objects have associated audio that is triggered when learner approaches the object. For instance, associated with the statue of Velázquez is a speech about his major paintings. Learners can also hear pre-recorded conversations between NPCs (see Fig. 1). Simple conversations allow illustrating the use of grammar patterns and more complex conversations, related to cultural aspects of the lesson topic, allow the development of more advanced listening skills.
- *Writing skills:* Learners will develop basic writing skills using the vocabulary and grammar of the lesson to ask and give information to NPCs that understand simple constructions. This is done by using natural language processing chatbots. For instance, at Fig. 2 David is asking a female chatbot for an address.
- *Speaking skills:* The activities previously described are achieved primarily through the exploration of the virtual environment. All the learners can discover the same vocabulary and language patterns, but not all of them

receive the same information about Velázquez. Learners are divided into groups and each group will hear only some of the dialogs played by the NPCs. Learners should exchange the information received among their pairs in order to pass collaboratively a final test. See the back of Fig. 2, where two learners are talking.



Figure 1. Practicing reading and listening skills

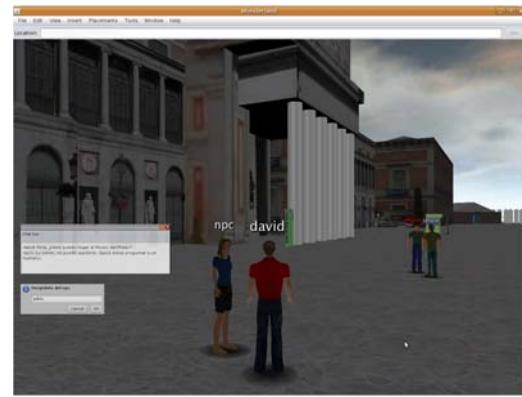


Figure 2. Practicing writing and speaking skills

Implications for an architecture to deploy 3DVWs to learn foreign languages

The deployment of any learning environment over 3DVWs, and in particular those based on situated learning, requires a 3D scenery filled with meaningful 3D elements: 3D objects relating to the context of the application, and NPCs to simulate real-life situations in the learning environment. It is expected that these graphical elements in a 3D medium provide visual immersion to learners and thus engage them in the learning experience. Visual immersion can be fostered by using hardware devices providing the stereoscopic vision, 360 degree immersive virtual reality.

Avatars are the mediums learners have to interact with the virtual world, communicate with other avatars, and navigate through the world. To cover the interaction capabilities, our application requires multimodal information attached to 3D elements that can be viewed or heard when a learner selects or approaches them. These scripting possibilities will support the reading and listening skills. A more sophisticated way of interaction that is especially useful in our application is through NPCs provided with Artificial Intelligence tools that allow them to understand simple written sentences. By including this capability, the application supports the development of writing skills.

None of the above is possible without an adequate means of navigation through the world. Usually this is done via the mouse or the keyboard. Nevertheless, it is also desirable to have software elements to overcome the orientation problems in 3D.

The collaboration activity designed to develop speaking skills, requires students to work in groups and the possibility to give different information to each group. Thus, the system must provide capabilities to group students and security mechanisms that allow restricted access to information.

Architecture

We have built our 3D virtual learning environment with Open Wonderland (Open Wonderland Foundation, 2010), a cross-platform, free and open source software. The toolkit is written entirely in Java, it supports audio conferencing, desktop application sharing, and integration with external data sources. This platform has been chosen by the “Immersive Education Initiative” (Immersive Education, 2009) to integrate an ecosystem of platforms in which learning objects can be exchanged.

Open Wonderland has a distributed client-server architecture. We have extended its functionality by plugging in several modules required in our learning environment (see Fig. 3). Although each module has three components

executed in the client, the server and both client and server respectively, a module is identified as a client's module when its primary functionality is executed on the client, otherwise is identified as a server's module.

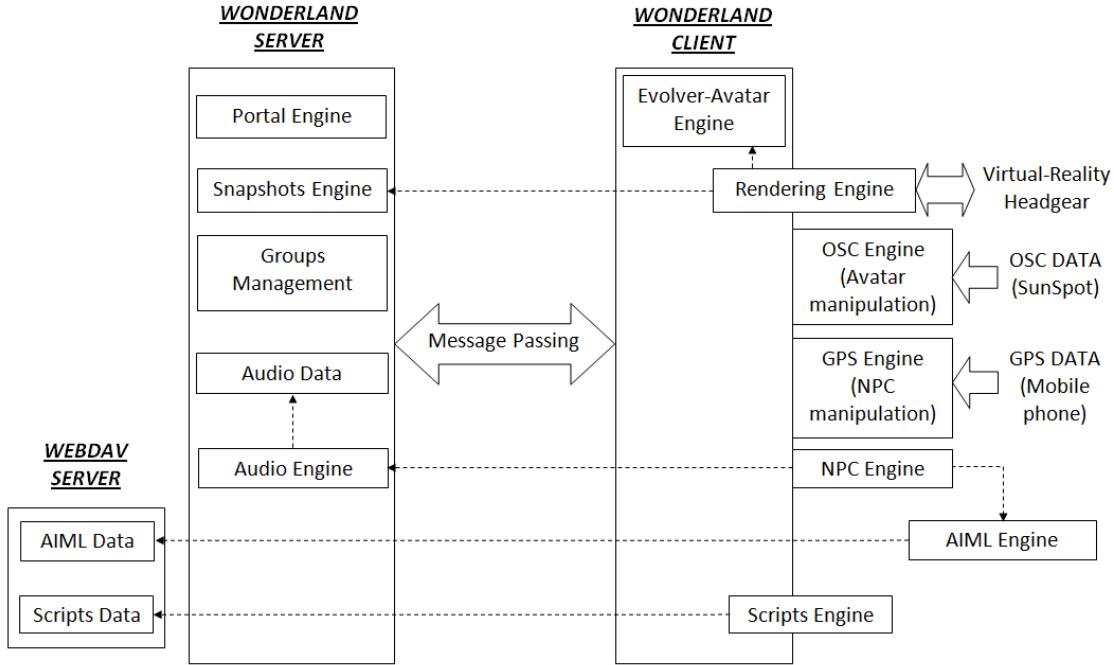


Figure 3: Architecture design

Open Wonderland Server

A 3DVWs is a composition of 3D scenes filled with NPCs, chatbots and smart objects that are installed by the administrator into the Snapshot Engine. In Open Wonderland, these graphical objects must be in the COLLADA (Arnaud and Barnes, 2006) format and are stored as XML files. In order to simplify the building of our 3D virtual learning worlds, we use the application Google SketchUp (Google, 2010) to create (or import) the 3D models required.

One advantage of Open Wonderland over other platforms is that it can be used to build collaborative 3D environments with spatial sound capabilities. These capabilities are provided by its Audio Engine and are particularly relevant for our Spanish learning environment because they provide full audio immersion. The audio immersion is achieved by attaching each audio treatment to a point in the 3D scene. The point can be, for instance, an NPC that identifies the source of audio, when an avatar approaches to that point, the user can hear the sound louder.

The audio data used to reproduce the NPCs conversations was obtained using a Text To Speech (TTS) technology that provides an acceptable quality. The audio files were created with TTS Reader (SpheNet, 2009) freeware software. As a future work, we are planning to use TTS technology to get the audio in real time, instead of using pre-recorded files.

A key aspect of our instructional design is the social interaction among students; this is implemented by grouping them in small units managed by the Group Management plug in.

Finally, each learning sequence must be orchestrated, it takes place in a scene and when the learning goals have been achieved, the avatar may be teleported to another learning sequence; this is done by the Portal Engine.

Open Wonderland Client

The visual component in a virtual world is essential to involve users in the virtual experience, therefore any platform to build 3DVWs provides a Rendering Engine to handle 3D graphics. In Open Wonderland 0.5 this engine is allocated in the client component and requires from the server's Snapshot Engine the XML files representing the virtual world. We extended Open Wonderland's Rendering Engine to provide visual immersion through the integration of a virtual-reality headgear as display. The new functionality allows the learner to switch into the full screen mode and set a 360° camera in order to watch through the world from any angle. The learner may also include new objects into the world; the easiest way to do so is by importing objects from Google SketchUp.

As avatars are users' representation in the virtual world, it is crucial that learners customize their avatars according to their preferences. As Open Wonderland 0.5 provides limited capabilities for this, thus we suggest using the Evolver 3D Avatar Generator (Darwin Dimensions, 2009) to create avatars and then import them into Open Wonderland's Avatar Engine.

To ease the movement of avatars in the virtual world, we have developed the OSC Engine, an avatar manipulation engine that allows students to move their avatars through a SunSPOT (Smith et al., 2006) (a video demonstration is available in <http://www.youtube.com/watch?v=kzd0AOHHIig>), besides the keyboard and the mouse. The SunSPOT communication with the Open Wonderland client is achieved with the OSC protocol (Wright, 2005), an Open Sound Control protocol optimized for modern networking technology.

Any virtual world platform provides mechanisms to add functionality to smart objects through Scripting Engines. Our platform customized programming object behaviors as reaction to mouse and key events. In this regard, the cursor's shape change when the mouse is over an object with information to the learner and text appears once the student clicks on a smart object. The changes made on the Scripting Engine promote reading skills and help students acquire vocabulary.

We distinguish among two types of characters controlled by our system: NPCs and chatbots. The first are synthetic characters which drive cyclical story line that perform dialogues depending on the student who approaches them. The last are used to transmit information to the students by simulating typical spanish people conversations. Chatbots encourage students to approach them when their avatars are in their surroundings, once done chatbots perform interactive dialogues with students. These behaviors contribute to the acquisition of listening and writing skills in our Spanish learning environment. ProgramD (Program D, 2009) an extended open source AIML platform was used to program the chatbots. As AIML (Artificial Intelligence Mark-up Language) is a XML-based programming language, it was necessary storing linguistic patterns and their possible answers related to the learning topic.

Finally, GPS Engine was developed to manipulate NPCs with an external device, a mobile phone with Symbian Operating System. The mobile phone uses the GPS technology to detect movement and send the NPC's new position to the GPS client's module via socket. In the future, we intend to use this technology to move the user's avatar.

WebDav Server

Data common to all clients are stored in a WebDav-based content repository hosted by the Open Wonderland Server. With this content repository, the client can access these data via the HTTP protocol.

AIML data and Script data needed by AIML Engine and Scripting Engine respectively are stored in the content repository. AIML data are the XML files that hold patterns that can be introduced by clients along with their associated answers. The Script data are JavaScript files holding the behavior associated with keyboard and mouse events.

Preliminary Evaluation of the Learning Environment

We conducted a preliminary evaluation to determine the usefulness of our learning environment in terms of motivation, immersion into a situated learning experience, and participation in collaborative activities.

The participants were twelve non-native Spanish speakers and six foreign language teachers grouped in six different experiences located in the virtual world. Only one participant had previous experience in virtual worlds. Participants did not have initial training using our system.

In the study, we observed the participants interacting with the 3DVW and we used interview evaluation techniques with open questions to identify strengths and weaknesses of the learning environment. Participants felt unsure when the tour began but they soon gained self-confidence. Collaboration emerged naturally to overcome initial difficulties participants had in understanding what to do and how to do it.

“At the beginning, I didn’t know where to go. I asked my friend and he told me what to do.”
“It is weird to walk through the middle of the street.”

Activities were perceived as games, participants were seen really engaged, and most of them continued the discussion after the experience had finished. In terms of communication, 3D audio provided a strong feeling of immersion, text chat was perceived as useful to establish communication with partners physically distant in the virtual world and most of users asked for tools for writing notes. Those who came from outside Madrid reported it was useful to be into a 3D scenery where the city’s cultural activity could be observed.

“Although I understood the dialogues, I would like to have a block of notes.”
“It was fun to chat with the actors (chatbots).”
“Are you sure that Velázquez was a friend of Quevedo?”
“Good, we are at Madrid!”
“I liked it, it is very visual.”

Despite features deployed to support orientation within the world, participants had difficulties to find locations. Furthermore, hardware devices provided to improve 3D visualization (Z800 3DVisor) and 3D interaction (SunSPOT), proved to be more a problem than a solution.

Conclusions and Future Work

3DVWs open the door to a new way of learning. Setting up realistic environments enhanced with a powerful set of learning oriented tools, these platforms allow for the implementation of sophisticated instructional models within a framework of richer information and cooperation.

In this paper, we have taken a step forward in the deployment of 3D virtual learning environments that fully exploit the immersive, interactive, and collaborative possibilities of 3DVWs. Technical and pedagogical features enrich our environments to provide students with formal and informal learning following less rigid curricula where a teacher is not always present. From the technical point of view, we included the use of haptic devices and natural chatting with NPCs. From the pedagogical point of view, we provide a collaborative environment where students will acquire and practice the necessary communication skills under the constructivist principles of situated learning and cooperative/collaborative learning.

We have conducted a preliminary evaluation to test the usefulness of our learning environment in terms of motivation, immersion into a situated learning experience, and participation in collaborative activities. The results were encouraging.

There is much more to improve in order to really convert a 3DVW environment into a learning platform. Another very important milestone will be the introduction of assessment procedures into 3DVWs, which is the challenge we are tackling now.

Acknowledgment

We thank the anonymous reviewers for their constructive criticism and suggestions for improving this paper.

This research is supported by the following projects: The Spanish CDTI project España Virtual within the Ingenio 2010 program. The Spanish project Learn3 (TIN2008-05163/TSI) within the Spanish “Plan Nacional de I+D+I”. The Madrid Regional Government project eMadrid (“Investigación y Desarrollo de tecnologías para el e-learning en la Comunidad de Madrid”), with grant No. S2009/TIC-1650.

References

- Andreas, K., Tsatsos, T., Terzidou, T., & Pomportsis, A. (2010). Fostering collaborative learning in Second Life: Metaphors and affordances. *Computers & Education*, 55(2), 603-615.
- Arnaud, R., & Barnes, M.C. (2006). *COLLADA: sailing the gulf of 3D digital content creation*. A K Peters, Ltd.
- Avatar Languages (2009). *Avatar Languages-Virtual learning brought to life*. Retrieved December 22, 2010, from <http://www.avatarlanguages.com>.
- Bell, M. (2008). Toward a Definition of “Virtual Worlds”. *Journal of Virtual World Research*, 1(1), Retrieved December 22, 2010, from <http://journals.tdl.org/jvwr/article/view/283/237>.
- Boulos, M.N.K.; Hetherington, L. & Wheeler, S. (2007). Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. *Health Information & Libraries Journal*, 24(4), 233–245.
- Callaghan, V., Gardner, M., Horan, B., Scott, J., Shen, L., & Wang, M. (2008). A Mixed Reality Teaching and Learning Environment. *Proceedings of the 1st international conference on Hybrid Learning and Education*, Berlin: Springer, 54-65.
- Calongne, C. M. (2008). Educational Frontiers: Learning in a VIRTUAL WORLD. *Educause Review*, 43(5), 36-48.
- Chittaro, L., & Ranon, R. (2007) Web3D Technologies in Learning, Education and Training: Motivations, Issues, Opportunities. *Computers & Education*, 49(1), 3-18.
- Clark, M.A. (2009). Genome Island: A Virtual Science Environment in Second Life. *Journal of Online Education*, 5 (6), Retrieved December 22, 2010, from http://innovateonline.info/pdf/vol5_issue6/Genome_Island-_A_Virtual_Science_Environment_in_Second_Life.pdf.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.
- Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10-32.
- Darwin Dimensions (2009). *Evolver*. Retrieved December 22, 2010, from <http://www.evolver.com/>.
- Dewey, J. (1916). *Democracy and Education*. New York: The Free Press.
- Dickey, M. D. (2005). Brave new (interactive) worlds: A review of the design affordances and constraints of two 3D virtual worlds as interactive learning environments. *Interactive Learning Environments*, 13(1-2), 121-137.
- Dillenbourg, P., Schneider, D., & Synteta, P. (2002). Virtual Learning Environments. *Communication*, 8(6), 3-18.
- Ertmer, P.A., & Newby, T.J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 6(4), 50-70.
- Eschenbrenner, B., Nah, F. F.-H., & Siau, K. (2008). 3-D Virtual Worlds in Education: Applications, Benefits, Issues, and Opportunities. *Journal of Database Management*, 19, 91-110.
- Genesee, F. (1985). Second Language Learning Through Immersion: A Review of U.S. Programs. *Review of Educational Research*, 55(4), 541-561.
- Girvan, C., & Savage, T. (2010). Identifying an appropriate pedagogy for virtual worlds: A Communal Constructivism case study. *Computers & Education*, 55(1), 342-349.
- Google (2010). *Google SketchUp*. Retrieved December 22, 2010, from <http://sketchup.google.com/>.
- Greenbush Education Service Center (2007). *Edusim – 3D virtual worlds for the classroom interactive whiteboard*. Retrieved December 22, 2010, from <http://edusim3d.com/>.
- Hendaoui, A., Limayem, M., & Thompson, C. W. (2008). 3D Social Virtual Worlds: Research Issues and Challenges. *IEEE Internet Computing*, 12(1), 88-92.
- Hinkel, E. (2006). Current Perspectives on Teaching the Four Skills. *Tesol Quarterly*, 40(1), 109-131.
- Holmes, J. (2007). Designing agents to support learning by explaining, *Computers & Education*, 48, 523-547.

- Huang, H. M., Rauch, U., & Liaw, S. S. (2010). Investigating learner's attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education*, 55 (3), 1171-1182.
- Immersive Education (2009). *Immersive Education Initiative*. Retrieved December 22, 2010, from <http://immersiveeducation.org/>.
- Instituto Cervantes. (2002). *Marco Común Europeo de Referencia para las Lenguas*. Assessment. Subdirección General de Cooperación Internacional.
- Jonassen, D. H. (1994). *Technology as cognitive tools: learners as designers*. Retrieved from http://aurorem.free.fr/partiels/sem7/cours/textesprincipaux/ITForum_Paper1_jonassen.pdf.
- Kalloni, P., & Sampson, D. (2010). Implementing a 3D Virtual Classroom Simulation for Teachers' Continuing Professional Development. *Proceedings of the 18th International Conference on Computers in Education* (ICCE 2010). Putrajaya, Malaysia.
- Kluge, S., & Riley, L. (2008). Teaching in Virtual Worlds: Opportunities and Challenges. *Issues in Informing Science and Information Technology*, 5, 127-135.
- Koenraad, T. (2008). *How Can 3D Virtual Worlds Contribute to Language Education?* Retrieved December 22, 2010, from <http://3dles.com/documents/worldcall2008-koenraad-revised2.pdf>.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, 19(3), 335-353.
- Lee, M. J. W. (2009). How Can 3d Virtual Worlds Be Used To Support Collaborative Learning? An Analysis of Cases From The Literature. *Society*, 5(1), 149-158.
- Linden Labs (2006). *New Media Consortium Campus*. Retrieved December 22, 2010, from <http://slurl.com/secondlife/NMC%20Campus/138/225/43>.
- Livingstone, D., & Kemp, J. (2008). Integrating Web-Based and 3D Learning Environments: Second Life Meets Moodle. *The European Journal for the Informatics Professional*, IX (3), 8-14.
- Nation, I.S.P., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking*, London: Routledge.
- Nieminen, K. (2006). Aspects of Learning Foreign Languages and Learning WITH Foreign Languages: Language Immersion and CLIL. *Development Project Report*, Jyvaskyla, Finland: Jyvaskyla University of Applied Sciences.
- Open Wonderland Foundation (2010). *Open Wonderland: Open source 3D virtual collaboration toolkit*. Retrieved December 22, 2010, from <http://openwonderland.org/>.
- Perkins, D. (1992). Technology meets constructivism: Do they make a marriage. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation*. (pp. 45-56). New Jersey: Lawrence Erlbaum Associates.
- Program D (2009). *Program D*. Retrieved December 22, 2010, from http://aitools.org/Program_D.
- Rauch, U., Cohodas, M., & Wang, T. (2009). The Arts 3D VLE Metaverse as a Network of Imagination. *Journal of Online Education*, 5(6), Retrieved December 22, 2010, from http://innovateonline.info/pdf/vol5_issue6/The_Arts_3D_VLE_Metaverse_as_a_Network_of_Imagination.pdf.
- Shih, Y.C., & Yang, M.T. (2008). A Collaborative Virtual Environment for Situated Language Learning Using VEC3D. *Educational Technology & Society*, 11(1), 56-68.
- Smith, R.B., Horan, B., Daniels, J., & Cleal, D. (2006). Programming the world with sun SPOTs. *Companion to the 21st ACM SIGPLAN symposium on Object-oriented programming systems, languages, and applications*, New York: ACM, 706-707.
- SpheNet (2009). TTSReader overview. Retrieved December 22, 2010, from <http://www.sphenet.com/TTSReader/index.html>.
- Three Immersions (2008). *3j School-Chinese Language*. Retrieved December 22, 2010, from <http://www.3jlife.com/cnSchool/index.jsp>.
- Vrasidas, C. (2000). Constructivism versus objectivism: Implications for interaction, course design and evaluation in distance education. *International Journal of Educational Telecommunications*, 6(4), 339-362.
- Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wilson, B. (1997). The postmodern paradigm. In C. R. Dills & A. Romiszowski (Eds.), *Instructional development paradigms*. Englewood Cliffs NJ: Educational Technology Publications, 297-309.
- Wright, M. (2005). Open Sound Control: an enabling technology for musical networking. *Organised Sound*, 10(3), 193-200.
- Zhang, Y. (2019). Cooperative Language Learning and Foreign Language Learning and Teaching. *Journal of Language Teaching and Research*, 1(1), 81-83.

Language Technologies to Support Formative Feedback

Adriana J. Berlanga¹, Marco Kalz¹, Slavi Stoyanov¹, Peter van Rosmalen¹, Alisdair Smithies² and Isobel Braidman²

¹Centre for Learning Sciences and Technologies (CELSTEC), Open University of the Netherlands, Heerlen, The Netherlands // ²University of Manchester Medical School, Manchester, United Kingdom // adriana.berlanga@ou.nl // marco.kalz.ou.nl // slavi.stoyanov@ou.nl // peter.vanrosmalen@ou.nl // alisdair.smithies@manchester.ac.uk // isobel.braidman@manchester.ac.uk

ABSTRACT

Formative feedback enables comparison to be made between a learner's current understanding and a desired learning goal. Obtaining this information is a time consuming task that most tutors cannot afford. We therefore wished to develop a support software tool, which provides tutors and learners with information that identifies a learner's progress, and requires only limited human intervention. The central idea is to use language technologies to create concepts maps automatically from texts, such as students' essays or Blogs. By comparing maps from students over time, or with maps created from tutor's materials, or by other students, it should be possible to ascertain learners' progress and identify remedial actions. We review existing tools for automatic construction of concepts maps and describe our initial explorations of one of these tools. This paper then introduces the theoretical background of the proposed tool, design considerations and requirements. An initial validation, which explored tutors' perceptions of the tool showed that tutors found the approach relevant, but its implementation in practice requires to consider teachers' practices, the tools already in use, as well as institutional policies.

Keywords

Formative feedback, Conceptual development, Concept maps, Language Technologies

Introduction

According to Hattie and Temperley (2007) effective feedback should provide information that helps students to see where they are going (learning goals); feedback information that tells students "how they are going", and feed forward information that points out to students "where to go next". From the tutor's perspective, providing this feedback requires several tasks, for example considering the learner's position regarding the curriculum (i.e., his/her current stage of learning), assessing his/her level of understanding, identifying possible gaps in knowledge, and suggesting remedial actions. These are time consuming tasks, especially as learners may have different learning goals and backgrounds, and may follow divergent learning paths.

We believe that providing this feedback should be part of the next generation of support and advice services needed to enhance individual and collaborative building of competences and knowledge creation. The premise is that language technologies, and particularly Latent Semantic Analysis (LSA) (Landauer, McNamara, Dennis, & Kintsch, 2007), could be used for this. LSA creates a mathematical model in which both the domain knowledge and the knowledge of the learner can be projected thereby enabling the progress of the learner to be analysed (Clariana & Wallace, 2007).

Our aim is to design a tool that provides learners and tutors information about a learner's conceptual development set side by side with the intended learning outcomes of the curriculum and of others in their learning group. The tool would use language technologies to extract such information automatically, enabling tutors to provide students with formative feedback in an efficient and time effective manner. This paper presents the design considerations and initial validation of such a tool. The first section presents the theoretical background, followed by design considerations and requirements. After this, the paper presents a review of existing technological solutions. We discuss the use of one of them when applied in a "mock-up" to explore the feasibility of our approach. Thereafter, the paper describes the initial validation of a first prototype of the anticipated service. It investigates the tutors' perceived relevance and satisfaction of the approach. Finally, the paper presents conclusions and future work.

Theoretical background

Feedback, is a key element in formative assessment, can be defined as information provided by an agent (e.g., teacher, peer, book, parent, tool) regarding aspects of performance or understanding (Hattie & Temperley, 2007) with the aim of modifying thinking in order to improve learning (Shute, 2008).

In contrast with summative assessment, formative assessment does not have the intention of summarizing or grading the achievement of a student for certification purposes (Sadler, 1989); it occurs typically after instruction and seeks to impact on learning, by providing knowledge and skills or to develop particular attitudes or to advise the student on learning strategies (Hattie & Temperley, 2007; Sadler, 1989). It could be used both by teachers and students. Formative feedback provides teachers with useful information for making decisions regarding delivery of a programme on the basis of students' progress diagnosis of any shortcomings in students' learning and its remediation (Shute, 2008). Students and teachers use formative feedback to monitor the strengths and weaknesses of students' performances, the former could be recognized and reinforced, whereas the latter modified or improved (Sadler, 1989). According to Shute (2008) formative feedback can reduce learners' uncertainty on how well they are performing, can reduce their cognitive load (particularly with novice or struggling learners), can potentially promote learning, and provide useful information for correcting misconceptions. Formative feedback strategies include providing learners with information that moves them forward in their conceptual development, empowering them as owners of their own learning as well as "instructional resources" for individuals (Black & Wiliam, 2009).

Our ambition is to design a formative feedback tool that, with minimal human intervention, provides tutors and learners with information about learners' conceptual development. The design considerations of the tool we envisage are grounded on three aspects: developing expertise, knowledge creation, and the process of measuring conceptual development. A full description is provided by Berlanga, Van Rosmalen, Boshuijzen, and Sloep (in press).

Briefly, it has been observed that in the development of expertise, novices and experts differ in the way they structure their utterances and knowledge. Novices do so in networks that are incomplete, loosely linked and solve problems in long chains of detailed reasoning steps throughout these networks. In contrast, experts have well structured and organized mental frameworks. They structure knowledge so problems may be solved by omitting reasoning steps rather than by proceeding one step at the time. Differences between novices and experts are closely reflected in the textual utterances expressed by novices in their evolving domain knowledge. Thus the way novices and experts express their use of concepts and how they relate them to one another's changes through time. This occurs in a systematically way and is based on learning experiences (Boshuizen & Schmidt, 1992; Arts et al., 2006; Boshuizen et al., 2004).

Second, theories of knowledge creation focus on how individuals and groups develop knowledge that is new to them. They stress that it is not transmitted untouched and unchanged from one knowledgeable person to another individual who is unknowing. In contrast, they emphasize that knowledge is constructed in a dialectical and social process. Not only explicitly stated knowledge and information is a source or result of this process but it also depends on a much bigger reservoir of tacit knowledge. Examples of knowledge creation theories include Stahl's knowledge building cycle (Stahl, 2006), and the "SECI" process of Nonaka et al. (2000). Stahl proposes a model in which individuals build their knowledge in a cycle which comprises personal understanding and collaborative knowledge, and assumes that the construction of knowledge is a social process. The SECI process describes the interplay between the individual and group learning as four connected and interacting processes of knowledge conversion: socialization, externalization, combination and internalization. These processes can take place in different levels of sophistication, depending on how people create and employ a context for implicit and explicit communication, the quality of the input in the process, etc.

The third consideration is related to the process of measuring conceptual development. If we aim to develop a tool that provides formative feedback, learners should be able to judge the quality of their work. To this end they need to (a) possess a concept of the standard (or goal or reference level) for which they are aiming; (b) compare their actual (or current) level of performance with the standard; and (c) engage in appropriate action which leads to some closure of the gap between them (Sadler, 1989).

A well-known example of the use of computer modelling techniques to approximate the structure of a metacognitive theory (Schraw & Moshman, 1995) is the structural approach proposed by Goldsmith et al. (1991) which analyzes

how an individual student organizes the concepts of a domain. This involves three steps: (1) eliciting the student's knowledge, (2) representing his/her elicited knowledge, and (3) evaluating this representation relative to some standard (e.g., reference model, an expert's organisation of the concepts in the domain, reference).

Design considerations and requirements

Based on the theoretical foundations discussed above, the design of the proposed tool is grounded in the idea that providing formative feedback should consider that:

- A learner's level of expertise is reflected in the way they use and relate concepts, when they express their knowledge;
- Learners develop their expertise in a knowledge building process, which encompasses cognitive and social perspectives; and,
- Learners should be provided with diverse ways of comparing their level of performance.

To this end, the service should provide learners with diverse ways of comparing their understanding against different reference models (Berlanga et al., 2009):

- Predefined reference model which considers intended learning outcomes described in, for instance, course material, tutor notes, curriculum information, etc.
- Group reference model, which considers the concepts and the relations that a relevant group of people (e.g., peers, participants, co-workers, etc.) used the most.

The idea is that the tool is used by a learner or a teacher to process text materials automatically (such as student input, learning materials, etc.), and in return obtain the most relevant concepts included in the input text and the relation between these concepts. The tool could then represent them visually as a concept map or as a list of concepts. If the text input consists of intended learning outcomes (such as course materials, books, etc.), then the result of the automatic process is the so-called predefined reference. If the text input consists of written output from a group of students (aggregated in a single text), then the tool produces the so-called group reference model.

The tool should also offer the possibility of generating comparisons between different texts inputs. For example, if a tutor decides to generate one concept map from a predefined reference model and another using text input from a student, they may be compared to identify which concepts the student is omitting from his/her text but are present in the learning materials. These comparisons could also be obtained from a group model and a predefined reference model. In this case the comparison will enable the tutor to identify those concepts that are not mentioned by the group as a whole and make it easier to identify outliers, diagnose causes of relevant problems, and take prompt remedial actions.

Based on our requirements and earlier work we decided to use language technologies as the underlying technology for the feedback tool. In order to decide how it could be implemented, next we review existing technologies that could support the analysis of conceptual development and serve as a foundation for this tool.

Existing tools for automatic construction of concept maps

In the previous section we have already referred to the use of concept maps. This means of eliciting and representing a learner's knowledge, is one of the most common ways of representing cognitive structures (Novak, 1998). Research evidence demonstrates that concept maps are well-suited for eliciting knowledge (Nesbit & Adesope, 2006), and are better for evaluating learners of different ages than classical assessment methods such as tests and essays (Jonassen, Reeves, Hong, Harvey, & Peter, 1997; Novak, 1998). The creation of concept maps, however, is a complex and time consuming task. It requires training and practice to understand how the relevant concepts should be identified and to make relationships between them. Therefore we analysed existing tools and tool sets that are able to support the creation of concept maps. We have not included purely algorithmic methods which have been tested for concept map construction (Bai & Chen, 2008, 2010) but we have focused on integrated solutions that allowed us to work with text input directly.

There are already a number of tools for the automatic construction and support of concept maps: Knowledge Network and Orientation (KNOT, PFNET) (Clariana, Koul, & Salehi, 2006); Surface, Matching and Deep Structure (SMD) (Ifenthaler & Seel, 2005); Model Inspection Trace of Concepts and Relations (MITOCAR) (Pirnay-Dummer, 2006); Dynamic Evaluation of Enhanced Problem Solving (DEEP) (Spector & Koszalka, 2004); jMap (Jeong, 2008) and Leximancer (Smith & Humphreys, 2006), Table 1 summarises these tools in terms of the data collection they require and the analysis and comparison they perform.

Table 1: Existing tools for construction of concept maps (adapted from Shute, Jeong, Spector, Seel, and Johnson, (2009))

| | Data Collection | Analysis | Comparison |
|------------|---|---|--|
| KNOT | Concept pairs/Propositions | Quantitative Analysis | Direct comparison of networks with some statistical results |
| SMD | Concept map or natural language | Quantitative analysis is calculated using tools | Unlimited comparison |
| MITOCAR | Natural language | Quantitative analysis included multiple calculations using tools | Paired comparisons for semantic and structural model distance measure |
| DEEP | Annotated causal maps | Quantitative/qualitative analysis is done mostly by hand | Unlimited comparisons, showing details relative to concepts |
| jMap | Concept maps, causal maps, or belief networks | Quantitative analysis (calculated using tools) | Superimposes maps of individual ($n=1$) and group of learners ($n = 2+$) over a specified target map |
| Leximancer | Concept maps | Content analysis and relational analysis (proximity, cognitive mapping) | Imposes tags in a single map over user-defined tags (names, concepts, files, etc.) |

These tools have some common characteristics: (a) they can (semi-)automatically construct concept maps from a text; (b) they use a type of distance matrix; (c) they propose a quantitative analysis of the maps; and (d) most of them are concerned with conceptual development of learners. Among their differences, we have found that, even though they all use a language technology, not all of them refer to it explicitly. These tools also differ on the scoring schemas they use to perform the quantitative analysis: DEEP uses the number of nodes and links; SMD uses propositions or the number of the links of the shortest path between the most distant nodes.

Most of these concept mapping tools provide opportunities to identify the conceptual gap between a learner's concept map and a criterion map (which could be a predefined reference model or group model), or to compare a learner's concept maps over different periods of time. However, only SMD, jMap and, in some extent DEEP, purposely provide a visualisation of this progression with reference to the standard criterion. Most of these mapping approaches construct and analyse individual maps. jMap visualises and assesses changes observed in either individual or collective maps. However, jMap is restricted to producing causal maps. KNOT, SMD, MITOCAR and Leximancer report on reliability and the correlation of validity criteria. Typically, they consist of the automatic scores generated by these tools, human concept mapping scores and human essay scores.

Each of the tools discussed can be used, at least to some extent, to provide formative feedback. Leximancer is the only tool that does not require specific input to start and/or a specific way of working. Therefore, based on our requirements we have focused on using Leximancer for an initial, empirical validation of our approach.

Initial explorations

In view of the theoretical considerations discussed above, we designed and prepared two experiments with functional mock-ups of the service to check the validity of our ideas. Each of the mock-ups was based on a combination of manual interventions and existing tools. The mock-ups were used to explore the following questions:

- A. Is it possible to build a concept map of a text on a selected topic that, according to the writer, covers the *core* concepts of the text?
- B. Similarly, is it possible to build a ‘group’ concept map which represents a set of selected texts on a specific topic that, according to the authors, covers the *core* concepts of the aggregated text?
- C. Do the writers of the input texts perceive the representations of A and B as useful input when they want to compare and contrast the individual versus the group perspective of the selected topic?

In the first experiment (Berlanga et al., 2009), users were only indirectly involved i.e., as providers of materials, as the actual outcomes were assessed by an expert. In the first test we transcribed a student’s spoken description of a medical case and used Leximancer to create a concept map (A) of this text and of the tutor materials for the corresponding topic (B). The results indicated that the student’s concept map used much more detailed concepts compared to that derived from the tutor materials. The study illustrated that a model based on comparing concept maps from tutor materials with those from students, must be used with care, since the interpretations of such maps may require more expertise than is possessed by a student in C, who is at a novice level.

In the second test of the first experiment, we used Leximancer to create concept maps (A) of each of 10 interviews with researchers in our group on how they understood the concept of a Learning Network (Sloep, 2008) and one emerging concept map based on integrated summary of all transcripts (B). Results indicated that by using Leximancer we identified the 10 most commonly used concepts and their importance. Moreover, an initial analysis showed that a comparison of an individual’s map and the group map could be used to indicate differences and similarities (C).

In the second experiment (Berlanga et al., in press), we explored the same questions. This time, however, the users were directly involved. We asked six researcher of our research group to provide us with one of their articles (average size 5000 words) on their research on Learning Networks. Each of the articles and the summary of all of them were represented as a concept map by Leximancer and, alternatively, as a word cloud by using Wordle (<http://www.wordle.net>) to check the possibilities of more commonly used tools. A questionnaire, based on the questions A, B and C, stated above, was used to assess the users’ perceptions of the concept map and the word cloud. The results indicated that there was a fair coverage of concepts included in the articles by both representations, in answer to question A. Likewise, in answer to B, the representations of the summary of all articles covered by the Learning Network were, as a whole, satisfactory. The answer to question C was more ambiguous; five of the six users, found the concept map was useful for detecting similar and missing concepts when their article was compared with the summary article, whereas three out of the six users obtained this results with the word cloud.

The results of the two experiments indicated that there were sufficient grounds to start developing a dedicated prototype.

Validation of the approach

Following the results from the partly manual explorations described above, the proposed design was used to develop a first prototype of an automated tool called CONSPECT (Wild, Haley, & Bülow, 2010). This tool enables a user to extract the core concepts from their own text and a reference text automatically. The comparison can be shown both as a list and a concept map (as shown in Fig. 2).

As a first step, the CONSPECT service was validated from the perspective of tutors at the University of Manchester (UK), who were involved in year 2 of a 5 year undergraduate medical degree. Five tutors were recruited for this purpose, four of whom had more than five years experience in this role and one was less experienced, but had been tutoring for over one year; all but one were women. The software was explained to all participants, who were given an overview and demonstration of CONSPECT and shown how to input materials and access outputs. The text

output used were blogs, written by students and tutors on the weekly clinical case studied in that part of the programme. They were trained to interpret results and asked to produce a 'model answer' blog. The concepts from the blogs were extracted by CONSPECT and were compared with those produced by the students, using either a student group reference model or blogs produced by individuals, which were also compared with each other. These comparisons were then shown to the tutors. A mixed methods approach was used to record and analyse their responses. They completed a questionnaire, comprised of forty three questions each with a five-point Likert scale, which covered aspects of time management, usability and efficacy of CONSPECT, and its role in augmenting teaching. Tutors then completed free text comments which were thematically analysed. The main findings were that tutors gave highest ratings to their knowledge and skills in using the software and to their efficient completion of tasks. Analysis of free text comments indicated that tutors appreciated the fundamental basis for CONSPECT and that it could provide rapid comparisons of students' understanding of the particular subject area with a "model" answer. It had the potential to identify those students who engaged at a more superficial level and others who might be delving more deeply into the subject matter, which might enable tutors to confirm those individuals who were outliers in their groups (Smithies, Braidman, Berlanga, Wild, & Haley, 2010).

A further validation was then conducted in the Open University of The Netherlands in the context of distance education by obtaining feedback from tutors about the relevance of using CONSPECT for their practice. The rest of this section describes this validation.

Method and data

Five tutors of the Open University of The Netherlands, Faculty of Psychology attended a workshop session, which included a demonstration, an individual hands-on session, a focus group discussion and completing a questionnaire.

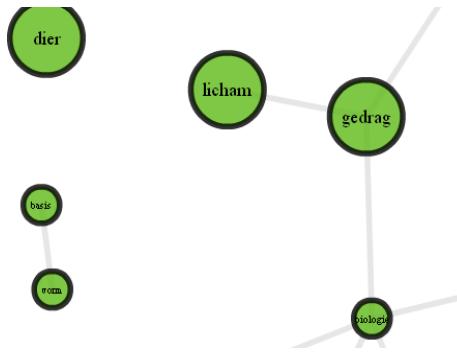


Figure 1a: Concept map from learning materials
(predefined reference model; zoom view)

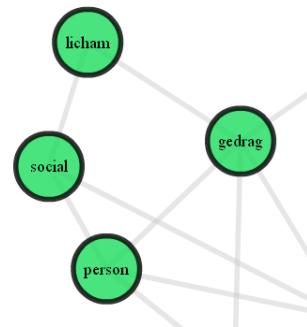


Figure 1b: Group concept map (zoom view)

In preparation for the workshop, we collected learning materials from two Psychology courses namely a digitalized book of the course, and a tutor's model answer for a specific assignment, which answered specific questions that covered the main course topics. We used CONSPECT to create a predefined reference model of the main concepts that the students should cover (see Fig.1a for an example). We also used examples from the students' answers to create a group reference model (Fig.1b). Finally, a comparison between these models was created alongside a list of similar and dissimilar concepts (Fig. 2), to identify concepts that are not covered well by the students.

During the first part of the workshop the design of the CONSPECT service and its aim were presented. Examples of concept maps were also introduced to show the type of information the service could provide. The participants then had a hands-on session in which they were asked to use the service with the existing materials to generate a concept map for a student, a concept map for a group model and a concept map for a predefined reference model. They were then required to compare these concept maps and see the results provided by the service. The respondents were asked to work alone, and take notes about their experience with the tool. If necessary, support was provided. Finally, we conducted a focus group, which was recorded both electronically and by notes taken at the time. Data was also collected as follows:

- Background questionnaire, to summarise tutor's teaching experience and age;

- Post activity questionnaire using a five-point Likert scale, with questions about relevance of the tool and user satisfaction:
 - Relevance: Four questions explored how tutors perceived the tool (see Table 2).
 - Satisfaction: Based on the UTAUT questionnaire (Venkatesh, Morris, Davis, & Davis, 2003), six questions were posed to explore the perceived satisfaction (see Table 3).
- Observer and participants' notes made during the validation session;
- Notes and audio recording from the focus group.

Participants (n=5) were tutors from different Psychology areas, with more than 5 years of experience in teaching. Three of them were between 30-40 years old; and the rest were older. Four of them were male.

| Concept map from Learning Material (see Fig. 1a) | | |
|---|--|--|
| Overlapping concepts: learning material and group map | | |
| Group concept map (see Fig. 1b) | | |
| erfelijk (hereditary) invloed (influence) kenmerk (characteristic) basis (basis) biologie (Biology) genetisch (genetic) sterk (strong) vorm (form, manner) | biologisch (biological) dier (animal, creature) gedrag (behaviour) licham (body) manier (way) men (people, one) mens (people) menselijk (human) ontwikkel (development) onderzoek (study) person (person) psychologie (Psychology) social (social) theorie (theory) verschill (difference) | natur (nature) ontwikkel (development) problem (problem) psycholog (psychology) wetenschap (science) |

Figure 2: Comparison between concept maps

Results

The results summarised in Table 2 show that all tutors considered the information provided by the tool is useful in identifying the progress of a group of learners (Q2), and that only 20% of them considered it was not useful for identifying the progress of individual learners (Q1). Not all tutors (40%) considered that the approach is relevant for addressing “burning” problems of their institution (Q3), whereas most (80%) of the tutors indicated that they could identify new potential uses of the tool (Q4).

Table 2: Perceived Relevance of the approach

| Question | Negative (<3) | Neutral (=3) | Positive (>3) |
|---|---------------|--------------|---------------|
| Q1. The way CONSPECT provides information (list concepts, graphical representation) is useful to identify learners progress | 20% | 60% | 20% |
| Q2. The way CONSPECT provides information (list concepts, graphical representation) is useful to identify the progress of a group of learners | - | 60% | 40% |
| Q3. I think CONSPECT addresses one of the burning problems of the institution | 40% | 20% | 40% |
| Q4. I can identify new potential uses of CONSPECT, which will address problems of the institution | 20% | 40% | 40% |

The results for tutor satisfaction are summarised in Table 3. Most tutors (80%) considered that the tool increases their curiosity about the topic (Q6), whereas only 40% indicated that the tool makes teaching more interesting (Q7); Half the tutors indicated that the tool motivates them to explore the teaching topic (Q8); and 40% considered themselves eager to explore the tool further (Q10). However, 60% of the tutors were negative regarding the way the tool would help them in their teaching (Q1), and on recommending the tool to other teachers (Q9).

Table 3: Satisfaction of the approach

| Question | Negative (<3) | Neutral (=3) | Positive (>3) |
|--|---------------|--------------|---------------|
| Q5. Overall, I am satisfied with the way CONSPECT would help me in my teaching. | 60% | 40% | - |
| Q6. Using CONSPECT increases my curiosity about the teaching topic. | 20% | - | 80% |
| Q7. CONSPECT makes teaching more interesting. | 60% | - | 40% |
| Q8. Using the CONSPECT motivates me to explore the teaching topic more fully. | 50% | - | 50% |
| Q9. I would recommend CONSPECT to other teachers to help them in their teaching. | 60% | 40% | - |
| Q10. I am eager to explore different things with CONSPECT | 60% | - | 40% |

The initial reaction of the tutors was positive, as they pointed out that indeed one of the problems they face is that they cannot easily identify students that are struggling with the course and that providing formative feedback promptly is a time consuming task. Tutors also feel students work only to get marks on assessments, instead of producing evidence of their actual learning. During the focus group, 4 out of 5 tutors commented that the approach has potential for their practice. They all stress, however, the importance of integrating the tool in their current learning environment, as essential for them to use the tool.

In their validation of the concept maps, tutors indicated that they could easily identify the most relevant concepts, as well as the similar and dissimilar concepts. They also mentioned that the maps had a fair coverage of the content and meaning of the text. Although tutors found it difficult to interpret the representation of the concept maps, they liked the idea of visualizing the links between the concepts, instead of simply a list of overlapping concepts.

The respondents felt that the user interface of the tool was still too complex for most people, but they acknowledged the added value of the approach. They suggested a variety of new forms in which the approach could be used in their teaching practice, for instance:

- Checking different resources (e.g., books, papers, articles), comparing them and deciding which is most relevant to the course learning objectives
- Checking if the learning materials produced by tutors contain the most relevant concepts
- Generating outlines (based on a set of input resources) to create study materials
- Initially checking the quality check of students texts, by asking them to write a text from which a concept map could be generated
- Using the tool in forums, to get a picture of what topics have been discussing
- Checking for plagiarism by comparing different student's texts.

Conclusions

In this paper we argued that a tool to provide prompt formative feedback can be designed in such a way that, by means of language technologies, little tutor intervention is needed. We proposed that learners will benefit if a tool provides them with information regarding their coverage of key concepts in the study domain, and compares this information with that of their peers. From the tutor perspective this feedback provides evidence which can then help identify individual learners who have difficulty in recognising key concepts.

From the validation conducted it was evident that most tutors perceived the approach relevant and useful for them and their students. They also suggested several different ways of using the tool, which indicated that they appreciated the potential of the tool. Nevertheless, tutors identified several conditions that should be fulfilled in order to incorporate the tool in their current practice. This may also negatively influence the results regarding user satisfaction. There were strong arguments in favour of aligning the tool with existing practices, such as total integration to existing platforms (e.g., institutional virtual learning environment), using only specific types of text documents, or privacy issues in sharing information. These constraints, whether institutional or tutor-oriented, are difficult to avoid if the proposed service is to be implemented in real practice. At the same time, these might cause that stakeholders overlook the potential technology has for supporting learning and therefore limit the possibilities the service –or any other new technology solution– could provide in the learning practice.

We believe that our approach could be of use in other learning situations, where different pedagogical approaches are used. It could be valuable in collaborative writing, where it is important to recognise differences and similarities between the texts, in discussion forums, to identify which concepts have been discussed, in workplace learning to specify core concepts in different documents (for a trial case see Berlanga et al. (2009)) and in informal learning situations where a formative feedback tool, such as the one we propose, could be of use to a group of people who share the same interest on a particular topic and are willing to explore the domain further.

Finally, further research is needed to evaluate learner's perception of the proposed tool as well as evaluation that involves a wider range of stakeholders. It is also essential to verify the accuracy and reliability of the language technologies used to underpin the development of this tool. This is important as we must ascertain how tutors and learners understand the limits of this technology, the conditions under which it may be used to produce reliable results, and those in which some results may be inaccurate.

Acknowledgments

We would like to thank Jannes Eshuis (Open University of The Netherlands), Jan Hensgens (Aurus KTS), Fridolin Wild and Debra Haley (Open University UK), and Robert Kobischke (Vienna University). We also like to thank the participants of the validation sessions. The work presented in this paper was carried out as part of the LTfLL project, which is funded by the European Commission (IST-2007-212578).

References

- Arts, A. J., Gijselaers, W. H., & Boshuijzen, H. (2006). Understanding managerial problem-solving, knowledge use and information processing: Investigating stages from school to the workplace. *Contemporary Educational Psychology*, 31(3), 387-410.
- Bai, S.-M. & Chen, S.-M. (2008). Automatically constructing concept maps based on fuzzy rules for adapting learning systems. *Expert Syst. Appl.*, 35(1-2), 41-49.
- Bai, S.-M., & Chen, S.-M. (2010). Using data mining techniques to automatically construct concept maps for adaptive learning systems. *Expert Systems with Applications*, 37(6), 4496-4503.
- Berlanga, A. J., Brouns, F., Van Rosmalen, P., Rajagopal, K., Kalz, M., & Stoyanov, S. (2009). Making Use of Language Technologies to Provide Formative Feedback. *Paper presented at the AIED 2009 Workshop Natural Language Processing in Support of Learning*, July, 6-7, Brighton, United Kingdom.
- Berlanga, A. J., Van Rosmalen, P., Boshuijzen, H. P. A., & Sloep, P. B. (in press). Exploring Formative Feedback on Textual Assignments with the Help of Automatically Created Visual Representations. *Journal of Computer Assisted Learning*.

- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Boshuizen, H. P. A., Bromme, R., & Gruber, H. (2004). On the long way from novice to expert and how travelling changes the traveller. In H. P. A. Boshuizen, R. Bromme & H. Gruber (Eds.), *Professional learning: Gaps and transitions on the way from novice to expert* (pp. 3–8). Dordrecht: Kluwer.
- Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153–184.
- Clariana, R., Koul, R., & Salehi, R. (2006). The criterion-related validity of a computer-based approach for scoring concept maps. *International Journal of Instructional Media*, 33(3), 317–325.
- Clariana, R., & Wallace, P. (2007). A Computer-Based Approach for Deriving and Measuring Individual and Team Knowledge Structure from Essay Questions. *Journal of Educational Computing Research*, 37(3), 211–227.
- Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83, 88–96.
- Hattie, J., & Temperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Ifenthaler, D., & Seel, N. M. (2005). The measurement of change: learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning*, 2(4), 317–336.
- Jeong, A. (2008). jMap v. 104, Retrieved May 1, 2011, from <http://dev22448-01.sp01.fsu.edu/ExcelTools/jmap/>.
- Jonassen, D., Reeves, T., Hong, N., Harvey, D., & Peter, K. (1997). Concept mapping as cognitive learning and assessment tools. *Journal of interactive learning research*, 8(3/4), 289–308.
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *Handbook of Latent Semantic Analysis*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Nesbit, J., & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76, 413–448.
- Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 33, 5–34.
- Novak, J. D. (1998). *Learning, creating and using knowledge: concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Erlbaum.
- Pirnay-Dummer, P. (2006). *Expertise and model building. MITOCAR*. Unpublished doctoral dissertation. University of Freiburg, Freiburg.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7, 351–371.
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189.
- Shute, V. J., Jeong, A. C., Spector, J. M., Seel, N. M., & Johnson, T. E. (2009). Model-Based Methods for Assessment, Learning, and Instruction: Innovative Educational Technology at Florida State University. In M. Orey (Ed.), *2009 Yearbook Educational Media and Technology*: Greenwood Publishing Group.
- Sloep, P. B. (2008). *Netwerken voor lerende professionals; hoe leren in netwerken kan bijdragen aan een leven lang leren. Inaugural address*. Open Universiteit Nederland. Available at <http://dspace.ou.nl/handle/1820/1559>. Heerlen.
- Smith, E., & Humphreys, M. S. (2006). Evaluation of Unsupervised Semantic Mapping of Natural Language with Leximancer Concept Mapping. *Behavior Research Methods*, 38(2), 262–279.
- Smithies, A., Braidman, I., Berlanga, A., Wild, F., & Haley, D. (2010). Using Language Technologies to support individual formative feedback. *Paper presented at the 9th European Conference on e-Learning*, November, 4-5, Oporto, Portugal.
- Spector, J. M., & Koszalka, T. A. (2004). The DEEP methodology for assessing learning in complex domains *Final report to the National Science Foundation Evaluative Research and Evaluation Capacity Building*. Syracuse, NY: Syracuse University.
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge: MIT Press.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: towards a unified view. *MIS Quarterly* 27(3), 425–478.
- Wild, F., Haley, D., & Bülow, K. (2010). Monitoring Conceptual Development with Text Mining Technologies: CONSPECT. *Paper presented at the EChallenges conference*, October, 27-29, Warsaw, Poland.

Designing for Automatic Affect Inference in Learning Environments

Shazia Afzal and Peter Robinson

University of Cambridge, Computer Laboratory, CB3 0FD United Kingdom //Shazia.Afzal@cl.cam.ac.uk //
Peter.Robinson@cl.cam.ac.uk

ABSTRACT

Emotions play a significant role in healthy cognitive functioning; they impact memory, attention, decision-making and attitude; and are therefore influential in learning and achievement. Consequently, affective diagnoses constitute an important aspect of human teacher-learner interactions motivating efforts to incorporate skills of affect perception within computer-based learning. This paper provides a discussion of the motivational and methodological issues involved in automatic affect inference in learning technologies. It draws on the recent surge of interest in studying emotions in learning, highlights available techniques for measuring emotions, and surveys recent efforts to automatically measure emotional experience in learning environments. Based on previous studies, six categories of pertinent affect states are identified; the visual modality for affect modelling is selected given the requirements of a viable measurement technique; and a bottom-up analysis approach based on context-relevant data is adopted. Finally, a dynamic emotion inference system that uses state of the art facial feature point tracking technology to encode the spatial and temporal signature of these affect states is described.

Keywords

Affective computing, emotions in learning, computer-based learning, facial affect analysis

Introduction

Computer-based learning now encompasses a wide array of innovative learning technologies including adaptive hypermedia systems to sophisticated tutoring environments, educational games, virtual environments and online tutorials. These continue to enrich the learning process in numerous ways. Keen to emulate the effectiveness of human tutors in the design and functioning of learning technologies, researchers have continually looked at the strategies of expert human teachers for motivation and are making directed efforts to make this machine-learner interaction more natural and instinctive. Detection of learners' affective states can give better insight into a learners' overall experience which can be helpful in adapting the tutorial interaction and strategy. Such a responsive interface can also alleviate fears of isolation in learners and facilitate learning at an optimal level. To enhance the motivational quality and engagement value of instructional content, affect recognition needs to be considered in light of its implications to learning technologies.

Effective tutoring by humans is an interactive yet guided process where learner engagement is constantly monitored to provide remedial feedback and to maximise the motivation to learn (Merill, Reiser, Trafton, & Ranney, 1992). Indeed, formative assessment and feedback is an important aspect of effectively designed learning environments and should occur continuously and unobtrusively as an integral part of the instruction (Bransford, Brown, & Cocking, 1999). In naturalistic settings, the availability of several channels of communication facilitates the constant monitoring necessary for such an interactive and flexible learning experience (Picard et al., 2004; de Vicente & Pain, 1998). One of the biggest challenges for computer tutors then is to achieve the mentoring capability of expert human teachers (van Vuuren, 2006). To give such a capability to a machine tutor entails giving it the ability to infer affect.

Learning has a strong affective quality that impacts overall performance, memory, attention, decision-making and attitude. Recent research provides compelling evidence to support the multiplicity and functional relevance of emotions for the situational and ontogenetic development of learners' interest, motivation, volition, and effort (Pekrun, 2005). It reflects the growing understanding of the centrality of emotion in the teaching-learning process and the fact that as yet this crucial link has not been addressed in machine-learner interactions (O'Regan, 2003).

Despite this recognition of affect as a vital component of learning processes and a context for cognition, computer-based learning environments have long ignored this aspect and have concentrated mostly in modelling the behaviour of a learner in response to a particular instructional strategy (Picard et al., 2004; du Boulay & Luckin, 2001). This relative bias towards the cognitive dimension of learning is now being criticised and the inextricable linkage between affective and cognitive functions is being stressed. This comes at a time when advances in the field of affective computing have opened the possibility of envisioning integrated architectures by allowing for formal representation,

detection, and analysis of affective phenomena. This increasing interest in building affect-sensitive human-computer interactions thus finds an important application in learning technologies (Cowie et al., 2001).

Building on a discussion of recent studies highlighting the relevance of emotions in learning, this paper describes different techniques for measuring emotions and efforts in automatic recognition and/or prediction of affect in learning contexts before proposing a parallel emotion inference system. This is not an exhaustive survey of the past work but a selected discussion of recent works highlighting the concern and those attempting to address it. Throughout this paper the terms ‘emotion’ and ‘affect’ will be used interchangeably.

Learning and Emotions

The neurobiology of emotions suggests that not only are learning, attention, memory, decision-making and social functioning affected by emotional processes but also that our repertoire of behavioural and cognitive options has an emotional basis. This relationship underscores the importance of the ability to perceive and incorporate social feedback in learning (Immordino-Yang & Damasio, 2007). Indeed, recent evidence from educational research supports the relationship of emotion with cognitive, motivational and behavioural processes (Pekrun, 2005; Turner, Husman, & Schallert, 2002). The seminal works of Boekaerts (2003), Pekrun, Goetz, Titz, and Perry (2002) and Meyer and Turner (2002) have pioneered the renewed surge of interest in affect and learning.

In a series of qualitative case-studies, Pekrun et al. (2002) explored the ‘occurrence and phenomenological structures of academic emotions’. They demonstrated that learners experience a rich diversity of positive and negative emotions; the most frequently reported being: anxiety, enjoyment, hope, pride, and relief, as well as anger, boredom and shame. Developing a multidimensional instrument, the Academic Emotions Questionnaire [AEQ], they conducted quantitative studies to test assumptions underlying Pekrun’s cognitive-motivational model (Pekrun, 1992). Using dimensions of valence (positive vs. negative) and activation (activating vs. deactivating) they distinguished four groups of emotions with reference to their performance effects – positive activating emotions (such as enjoyment of learning, hope, or pride); positive deactivating emotions (e.g., relief, relaxation after success, contentment); negative activating emotions (such as anger, anxiety, and shame); and negative deactivating emotions (e.g., boredom, hopelessness). Accordingly, they studied the effects of these emotions on learning and achievement with cognitive and motivational mechanisms like motivation to learn, strategies of learning, cognitive resources, and self-regulation. Instances of these mechanisms like interest and effort, learning strategies like elaboration or rehearsal, task irrelevant thinking diverting cognitive resources and self-regulated learning as compared to reliance on external guidance may all occur in the course of learning with a computer tutor and are thus directly relevant to this study.

To evaluate the dynamic and interactive effects of affect and motivation on learning processes like task engagement and appraisal, Boekaerts (2003) conducted several longitudinal studies using the On-line Motivation Questionnaire (Boekaerts, 2002) and found evidence for the existence as well as relevance of two separate, parallel processing pathways – *the cold cognition pathway* and *the hot cognition pathway*. The cold cognition pathway consists of meaning-generating processes that are the building blocks of learning comprehension and problem-solving. The hot cognition pathway on the other hand comprises of the emotional evaluations of learning opportunities that are triggered by emotions and moods in the actual learning episode. In her Model of Adaptive Learning (Boekaerts, 1992), these represent the mastery and the well-being path respectively. Boekaerts asserts that the evaluative information of the hot cognition path is situation specific and initiates concern-related monitoring, thereby influencing both decision-making (short-term effect) as well as value attribution (long-term effect).

Based on a decade of research on motivation and a diverse study of learner-teacher interactions, Meyer and Turner (2002) highlight the inseparability of emotion, motivation and cognition; and argue for integrated approaches to treat these as equal components in the social process of learning. They report their findings as *serendipitous*, thus emphasising the presence of emotion in instructional interactions. Although the context of their research is classroom based, they provide a reflective account on the obvious nature of emotion in learning interaction.

Kort, Reilly and Picard (2001) highlight the importance of continuous affect monitoring as a critical mentoring skill. They propose a spiral model that combines the phases of learning to emotion axes by charting out quadrants that map different stages occurring in the learning process. The horizontal emotion axes range from negative to positive across

different emotion sets like anxiety-confidence, boredom-fascination, frustration-euphoria, dispirited-encouraged and terror-enchantment. The vertical axis forms the learning axis that represents the transition between *constructive learning* and *un-learning*. This model assumes that the learning experience involves a range of emotions in the space of the learning task and visualises the movement of a learner from one quadrant to another.

In an attempt to understand the emotional dimension of online learning in qualitative terms, O'Regan (2003) explored the *lived experience* of students taking online learning courses. The study identifies both positive and negative emotions experienced by students, significantly - frustration, fear/anxiety, shame/embarrassment, enthusiasm/excitement and pride. These had a variable effect on the learning process depending on the strength and nature of the emotion, as well as the associated learning context. In another study, using a manual affect coding system, Craig, Graesser, Sullins, and Gholson (2004) observed the occurrence of six affect states during learning with an intelligent tutoring system. They analysed frustration, boredom, flow, confusion, *eureka* and neutral, and found significant relationships between learning and the affective states of boredom, flow and confusion.

More recently, Jarvenoja and Jarvela (2005) and Wosnitza and Volet (2005) provide empirical evidence from participants in social online learning to categorise sources of emotional experience along self, task, context or social directedness to highlight the impact of students' emotions on their motivation and engagement in the learning process.

In essence, learning has a strong affective quality that impacts overall performance, memory, attention, decision-making and attitude (Kort, Reilly, & Picard, 2001; Lisetti & Schiano, 2000). We know from a multitude of studies in different educational contexts that learners experience a wide range of positive and negative emotions. These emotions are situated and have social and instructional antecedents. For the discourse to be effective, it is imperative then to have access to and ensure the emotional well-being of learners. Since learning with computers is essentially self-paced, assessing the learner's experience becomes important. The aim is to reasonably emulate the social dynamics of human teacher-learner interactions in models that capture the essence of effective learning strategies like one to one tutoring (Bloom, 1984; van Vuuren, 2006).

Measuring Emotions

Current methods for measuring emotions can be broadly categorised as Subjective/Objective and Qualitative/Quantitative. In the context of learning, an additional categorisation as Snapshot/Continuous can be defined based on the timing of the emotion measurement (Wosnitza & Volet, 2005). Snapshot type measurements are done immediately before/after the learning process while continuous measurements are process-oriented and give access to the ongoing emotional experience. Consequently, snapshot measures provide only a limited window into the anticipated or reflected emotions at the end of the learning experience as against the continuous measures that provide direct access to emotions as they unfold during learning. Table 1 categorises some common methods for measuring emotional experience during learning.

Table 1. Methods for measuring emotional experience during learning

| | Snapshot Type (Before / After Learning) | | Continuous Type (During Learning) | |
|------------|--|--|--------------------------------------|---|
| | Qualitative | Quantitative | Qualitative | Quantitative |
| Subjective | Open Interviews Emotional Probes Stimulated Recall | Questionnaires Surveys | Emotional Diaries Think-aloud | Experience / Time-Sampling |
| Objective | Structured Interviews | Transcripts Analysis Video Analysis | Observational Analysis | Interactional Content Physiology / Nonverbal Behaviour Analysis |

For intervention to be effective, remedial action has to be appropriately timed - particularly in the case of strong emotions. Given the complex and transient nature of emotions, any retrospective accounts are problematic because of issues related to the potential for multiple levels of awareness, reappraisals and reconstruction of meanings during recall (Schutz, Hong, Cross, & Obson, 2006). This necessitates dynamic evaluation of emotions but without disrupting the learning task itself. Ideally then, an unobtrusive, quantitative, and continuous account of emotional

experience is a suitable method of enquiry. Amongst the methods listed in Table 1, analysis of nonverbal behaviour in the lower right quadrant, offers a reasonable fit to this requirement (Pekrun, 2005; Picard, et al., 2004; Hudlicka, 2003). Analyses of tutoring sessions have indeed revealed that affective diagnoses, as an important aspect of expert human mentoring, depend heavily on inferences drawn from facial expressions, body language, intonation, and paralinguistic cues (Lepper, Woolverton, Mumme, & Gurtner, 1993). Advances in the field of affective computing have opened the possibility of emotion recognition from its nonverbal manifestations like facial expressions, head pose, body gestures, voice and physiology. The field is promising, yet in a formative stage as current technologies need to be validated for reliability outside controlled experimental conditions.

Automatic Measurement of Affect

The semantics and manifestation of affective phenomena have been extensively studied across the disciplines of psychology, cognitive science, computer vision, physiology, behavioural psychology, etc. In spite of this, it still remains a challenging task to develop reliable affect recognition technologies. The reasons are varied. Expression and measurement of affect, and specifically its interpretation, is person, time and context dependent. Sensory data is ambiguous and incomplete as there are no clear criterions to map observations onto specific affect states. Lack of such ground-truths makes validation of developed techniques difficult and worse still, application-specific. Consequently, we do not know whether a system that achieves higher classification accuracy than another is actually better in practice (Pantic & Rothkrantz, 2003). Affect modelling in real-time is thus a challenging task given the complexity of emotions, their personal and subjective nature, the variability of their expression across, and even within, individuals, and frequently, lack of sufficient differentiation among associated visible and measurable signals (Hudlicka, 2003).

However, despite the difficulties, a whole body of research is persevering to give computers at least as much ability as humans have in recognising and interpreting affective phenomena that enables them to carry out intelligent behaviour and dialogue with others. This optimistic vision has already produced some commendable results and the following section reviews how machine perception of affect is being realised within learning environments. The interested reader is referred to Zeng, Pantic, Roisman, and Huang (2009) for a survey of general affect recognition methods using audio-visual modalities.

Prior Work

Despite the prospects, there are relatively few studies on automatic affect sensing in learning environments. Table 2 compares these in chronological order based on the affect construct they measure, the information source they use, the learning context in which the study was done, and the specific computational approach adopted. Most of the works reviewed here measure discrete emotion categories like confusion, interest, boredom, etc. (Mavrikis, Maciocia, & Lee, 2007; Kapoor & Picard, 2005; D'Mello, Picard, & Graesser, 2007; and Sarrafzadeh, Fan, Dadgostar, Alexander, & Messom, 2004); while a few use appraisal-based models of emotion (Jaques & Vicari, 2007; Heylen, Ghijssen, Nijholt, & Akker, 2005; Conati, 2002). Related constructs like difficulty, stress, fatigue and motivation have also received some attention (Whitehall, Bartlett, & Movellan, 2008; Liao W, Zhang, Zhu, Ji, & Gray, 2006; de Vicente & Pain, 1998).

Based on the modelling approach used, affect inference methods can be broadly categorised as (Liao et al., 2006; Alexander, Hill, & Sarrafzadeh, 2005):

- Predictive - those that predict emotions based on an understanding of their causes
- Diagnostic - those that detect emotions based upon their physical effects, and
- Hybrid - those that combine causal and diagnostic approaches

Table 2. Affect modelling in learning environments

| Citation | Affect Construct | Information Source | Learning Context | Method |
|--|--|--------------------|------------------|---|
| Whitehall, Bartlett & Movellan (2008) | Difficulty level and speed of content | Facial expressions | Lecture videos | Support Vector Machines and Gabor filters |

| | | | | |
|---|--|--|--|---|
| Zakharov, Mitrovic & Johnston (2008) | Positive and negative valence | Facial expressions | Pedagogical agent-based educational environment | Rule-based system |
| Baker (2007) | Off task behaviour | Interaction Log files | Cognitive Tutor software | Latent response model |
| Jaques & Vicari (2007) | OCC Cognitive Theory of Emotions | User's actions & interaction patterns | Pedagogical agent based educational environment | Belief-Desire- Intention (BDI) reasoning; appraisal based inference |
| D'Mello, Picard & Graesser (2007) | Flow, confusion, boredom, frustration, eureka & neutral | Posture, dialogue and task information | Dialogue based ITS- Auto Tutor | Comparison of multiple classifiers |
| Kapoor, Burleson & Picard (2007) | Pre-frustration & Not pre-frustration | Facial expressions, posture, mouse pressure, skin conductance, task state | Automated Learning Companion | Gaussian process classification; Bayesian inference |
| Mavrikis, Maciocia & Lee (2007) | Frustration, confusion, boredom, confidence, interest & effort | Interaction logs & situational factors | Interactive Learning Environment- WALLIS | Rule induction |
| Liao et al. (2006) | Stress & fatigue | Physical appearance, physiological, behavioural and performance measures | Maths and audio based experimental tasks | Influence Diagram; Ensemble of classifiers |
| Amershi, Conati & Maclarens (2006) | Affective reactions to game events | Skin conductance, heart rate, EMG | Educational game- Prime Climb | Unsupervised clustering |
| Kapoor & Picard (2005) | Interest, Disinterest, break-taking behaviour | Facial expressions, posture patterns & task state | Educational Puzzle | Ensemble of classifiers |
| Heylen et al (2005) | Scherer's Component Process Model | Facial Expressions, task state | Agent-based ITS for nurse education- INES | Appraisal using stimulus evaluation checks |
| Sarrafzadeh et al (2004) | Happiness/success surprise/happiness sadness/disappointment, confusion frustration/anger | Facial expressions | Elementary Maths ITS | Fuzzy-rule based classification |
| Litman & Forbes (2003) | Negative, neutral & positive emotions | Acoustic-prosodic cues, discourse markers | Physics Intelligent Tutoring Spoken Dialogue System – ITSPOKE | Comparison of multiple classifiers |
| Conati (2002); Conati & Zhou (2004) | OCC Cognitive Theory of Emotions | Interaction patterns, personality, goals | Educational game- Prime Climb | Dynamic decision network; Appraisal based inference |
| de Vicente & Pain (2002; 1998) | Motivation | User actions and interaction patterns; Experience sampling | Japanese numbers ITS-MOODS | Motivation Diagnosis Rules |

The predictive approach takes a top-down causal view to reason from direct input behaviour like state knowledge, self-reports, navigation patterns or outcomes to actions. It is generally based on sound psychological theories like Scherer's Component Process Model (Scherer, 2005) or the OCC Cognitive Theory of Emotions (Ortony, Clore, &

Collins, 1998). The appraisal theory provides a detailed specification of appraisal dimensions along emotion-antecedent events like novelty, pleasantness, goal-relevance, coping potential and norm/self compatibility; but suffers from the methodological problem of reliance on an accurate self-appraisal. The OCC theory on the other hand defines 22 emotions arising as valenced reactions to situations consisting of events, actors and objects. It does not however include some important affect states like boredom, interest and surprise which are relevant to learning scenarios (Picard, et al., 2004).

Conati (2002) and Conati and Zhou (2002) implement the OCC theory to assess learner emotions during interaction with an educational game. They use a dynamic decision network to model affect states but do not establish the accuracy of the model empirically. In another study, de Vicente and Pain (2002) were able to formalise inference rules for diagnosis of motivation using screen capture of learner interactions with a tutoring system. This work is significant in that it relies only on the concrete aspects of learner interactions such as mouse movements and quality of performance for motivation inference. These rules however, have not been implemented and hence remain a theoretical assumption. Heylen et al. (2005) describe an attempt to relate facial expressions, tutoring situation and the mental state of a student interacting with an intelligent tutoring system. They do not infer affect states automatically from facial expressions but use Scherer's Component Process Model (2005) of emotion appraisal using stimulus evaluation checks. Their results are inconclusive and specific to the tutoring system used in their study.

Diagnostic methods on the other hand take a bottom-up approach and are based on inference from sensory channels using traditional pattern classification techniques to approximate or estimate affective behaviour. These rely on the understanding that non-verbal behaviour through bodily gestures, facial expressions, voice, etc, is instinctively more resourceful and aims to infer affective cues with the aid of sensors. Notable in this category is the Affective Computing Group at MIT which is involved in a series of projects towards the building of a Learning Companion. Kapoor et al. (2007) use a novel method of self-labelling to automatically classify data observed through a combination of sensors, into 'pre-frustration' or 'not-pre-frustration'. In related work, Kapoor and Picard (2005) use multi-sensor classification to detect interest in children solving a puzzle by utilising information from the face, posture and current task of the subjects. The high recognition rates on these classification techniques are achieved for a single distinct affect state using sophisticated and fragile equipment. These do not as yet perform real-time classification.

D'Mello and Graesser (2007) use posture patterns along with dialogue, to discriminate between affect states during interaction with an intelligent tutoring system called Auto-Tutor. This is a dialogue based system achieving recognition of affect states like flow, confusion, boredom, frustration, eureka and neutral. Interestingly however, the ground truth used for validating their classification is mainly the facial action coding of recorded interaction by FACS experts. FACS or the Facial Action Coding System is the anatomic classification devised by Ekman and Friesen (1978) that defines 44 Action Units to describe any human facial expression.

Amershi et al. (2006) use unsupervised clustering to analyse students' biometric expressions of affect that occur within an educational game. Their approach is quite interesting and different from the usual supervised classification techniques normally applied for automatic sensing. However, lack of a benchmark or standard to compare performance makes it difficult to evaluate the efficiency of this method.

Sarrafzadeh et al. (2004) employ a fuzzy approach to analyse facial expressions for detecting a combination of states like happiness/success, surprise/happiness, sadness/disappointment, confusion and frustration/anger. They do not, however, give a measure of the accuracy of their method and focus more on the stage after detection. Litman and Forbes (2003) propose a method of affect modelling from acoustic and prosodic elements of student speech. Their study is particularly relevant for dialogue based systems.

Recent works of Zakharov, Mitrovic, and Johnston (2008) and Whitehall, Bartlett, and Movellan (2008) that use facial expression analysis techniques to measure valence and difficulty level, respectively, also fall within this category.

Finally, models of hybrid approaches, as in Conati (2002) and Liao et al. (2006), leverage the top-down and bottom-up evidence in an integrated manner for improved recognition accuracy. This involves using dynamic probabilistic approaches to model uncertainty in affect and its measurement, while explicitly modelling the temporal evolution of emotional states. Such frameworks are promising as they can allow context-sensitive interpretation of affective cues.

However, specification and fusion of information from the multiple channels still remains a significant challenge for actual implementation.

Discussion and Scope of this Work

Ideally, automatic sensing should be able to function in real-time; measure multiple and co-occurring emotions unobtrusively and without causing disruption in the actual learning process. As reviewed in the previous section, numerous efforts are being made towards this goal to give computer-based tutoring some semblance of emotional intelligence. Table 2 lists the relevant works and categorises these according to their specific focus and approach. It highlights the variety in modelling techniques that range from rule-based systems to complex probabilistic models; the different ways in which affect is conceptualised in these systems based on whether a dimensional, discrete or appraisal-based stance is adopted; the array of interactional as well as behavioural measures used to infer affect; and importantly, the nature and focus of the learning setup used. Given this diversity in the measured affect constructs, the specific learning environments and the channels used as information sources; it is difficult to comment on the overall performance of a system and determine its efficiency in a broad sense. This inability to make generalisable claims is an acknowledged limitation of affect sensing technologies (Pantic & Rothkrantz, 2003) and makes it challenging to establish the merit and success of a particular system satisfactorily and with confidence. Nevertheless, what is apparent is a growing understanding of the importance of affect modelling in learning and this substantiates further research in the area. The following sections lay out some design choices that set the scope of this work and therefore the proposed system.

Conceptualisation of affect

The issue of representation is at the core of emotion research and therefore affective computing. This is because handling of emotion data by machines requires programmed representations of affect and a clear structure that will perform real-time interaction with a user. Selection of an appropriate descriptive framework embodies the way affect is conceptualised within a system, the way it is observed and assessed, and consequently, the way it is processed (Peter & Herbon, 2006). However, the question of representation is not a simple one as it requires an understanding of the typology and semantics of the whole range of emotion-related phenomena like short-lived, intense emotions; moods; long-lasting established emotions; stances; attitudes/preferences, traits/affect dispositions, etc (Cowie & Cornelius, 2003). All this complicates the task of describing emotional content and while no single best representation scheme exists, there are established psychological traditions that have been used effectively to formalise the behaviour of interest. One of most long-standing way by which affect has been described by psychologists is in terms of discrete categories – an approach rooted in everyday language and driven by historical tradition around the existence of universal emotions.

The main advantage of the categorical scheme is that people use it to describe emotional displays in everyday interactions and is therefore intuitive. However, assignment of emotions into discrete categories or words is often considered arbitrary because of the social and cultural differences in semantic descriptions of emotion and for a designer of an HCI system, the requirement of an exclusive unambiguous representation. Linguistic labels can be imprecise and capture only a specific aspect of the phenomena with an associated uncertainty in the perceived meaning of a category. Nevertheless, this approach has had a dominating influence on the field of affective computing and most of the existing systems focus on recognising a list of basic emotions. Traditional psychological lists of emotions are mostly oriented to archetypal emotions and these are not the states that appear in most naturalistic data, especially in HCI contexts. As such, they do not represent the full range of emotions that can occur in natural communication settings. To overcome the intractable number of emotion terms and to ensure relevance in potential applications, the strategy of preselecting context-relevant word lists or cumulating relevant categories to derive pragmatic lists as in the HUMAINE database (Douglas-Cowie et al., 2007) or the more principled taxonomy of complex mental states by Baron-Cohen (2004), has been advocated and applied effectively (Cowie, 2009; Zeng et al. 2009).

Following such an application-oriented approach, we considered emotion groups of annoyed, anxious, bored, confused, happy, interested, neutral and surprised using the taxonomy of complex mental states by Baron-Cohen (2004). This is a lexical taxonomy that groups together semantically similar emotion concepts so that each group

encompasses the finer shades of an emotion concept. Confusion for example includes states like unsure, puzzled, baffled and clueless while Happy includes pleased, cheerful, relaxed, calm, enjoying, etc. These encompass representative emotions from each of Kort, Reilly and Picard's (2001) emotion axes as well as those of Pekrun et al.'s (2002) academic emotions with the exception of hope, pride and shame which have more complex social antecedents and meanings and are therefore excluded from this study. The selected emotion descriptors thus have a wider scope than considered by previous methods.

Choice of Modality

Emotion is expressed through visual, vocal and physiological channels. The visual channel includes facial expressions, body gestures, eye-gaze and head pose; the vocal channel focuses on measures of intonation and prosody; while the physiological channel includes measures of skin conductance, blood volume pressure, heart rate, temperature, etc. Lack of a consistent mapping between observable aspects of behaviour and actual affective states, technical feasibility, and practical issues complicate the choice of modality for sensing in a learning setting. Issues of ethics, privacy and comfort further constrain the design, use and deployment of appropriate sensing technologies. The use of physiological sensing in particular is challenging. Though relatively easy to detect and reasonably unobtrusive now, physiological sensing has some inherent shortcomings like requirement of specialised equipment, controlled conditions, baseline determination and normalising procedures, possible discomfort in usage, expertise in use of sensing apparatus and issues of privacy and comfort (Scherer, 2005; Hudlicka, 2003). Speech analysis may not always be suitable as not all learning environments are dialogue based. Table 3 below gives a brief comparative overview.

Table 3. Overview of the three dominant channels of nonverbal behavior

| Visual | Vocal | Physiological |
|---|--|---|
| Facial expressions, Head pose, Body gestures, Eye-gaze | Speech, Prosody and Intonation | Skin conductance, Blood volume pressure, Heart rate, Breathing rate, Temperature, Muscle tension |
| <ul style="list-style-type: none"> • Natural and observable • Unobtrusive • Practically deployable • Does not require specialised equipment; exception for gestures and eye-gaze • Behavioural coding required to set ground-truth | <ul style="list-style-type: none"> • Natural, discernable • Unobtrusive • Practically deployable • Limited to dialogue based systems • Manual annotation required to set ground-truth | <ul style="list-style-type: none"> • Unobservable • Unobtrusive but has issues with comfort and privacy • Requires tightly controlled environmental conditions • Specialised and fragile equipment • Easy to access the bio-signals but difficult to interpret |

As reviewed in previous works listed in Table 2, multiple channels are currently being probed for emotional signs ranging from facial expressions, posture, pressure patterns, prosody, interaction patterns and even trait factors like personality. Combination of one or more channels is likely to improve accuracy of emotion but is a challenging problem and a research avenue in itself. An important issue here is to understand redundancy and variation in the time course of the different information channels to inform purposeful fusion of relevant information. Works like that of D'Mello, Picard, and Graesser (2007) who analyse relative contributions of information channels are important for viable design and implementation of such systems.

Given the pre-eminence of facial signs in human communication the face is a natural choice for inferring affective states. With the latest computer vision techniques facial information can be detected and analyzed unobtrusively and automatically in real-time requiring no specialized equipment except a simple video capture device. This makes facial affect analysis an attractive choice for evaluating learner states and together with head gestures is selected as the modality for affect inference in our system. Moreover, although recent studies have looked at the divergence in emotional information across modalities (Cowie, 2009; Cowie & McKeown, 2009), affect inference from facial expressions has been found to be consistent with other indicators of emotion (Cohn, 2006). However, facial expressions are not simple read-outs of mental states and their interpretation being context-driven is largely situational. Computer tutors can exploit this aspect to infer affective states from observed facial expressions using the knowledge state and navigation patterns from the learning situation as supporting evidence. Given the requirements of an affective computer tutor, the visual modality thus has a great potential for evaluating learner states thereby

facilitating an engaging and optimal learning experience. It is for these reasons that the visual modality was selected for affect analysis in this work.

Context and Corpora

For a meaningful interpretation and to ensure ecological validity, it is essential to study the occurrence and nature of affect displays *in situ*, as they occur. Although a number of face databases exist, these are mostly posed or recorded in scripted situations that may not be entirely relevant in a learning situation. We know that emotions are situated, have contextual antecedents and are influenced by social consequences. Knowledge of the learning setting is important then to ground a research work in a specific context and help assess its generalisation ability. The nature and dynamics of emotions in a solo learning setting e.g., Conati and Zhou (2002), Conati (2002), will no doubt differ from those generated within an agent-based learning environment like in Jaques and Vicari (2007), Heylen, Ghijssen, Nijholt, and Akker (2005), Kapoor, Burleson, and Picard (2007), or with those that involve dialogue, as in D'Mello, Picard, & Graesser (2007), Litman and Forbes (2003).

The nature of affect and its dependence on context thus makes the choice of a learning environment an important one. As such, we decided to use a solo, one to one learning setting for our study. By focusing on a self-regulated learning model our objective was to minimise the potential effects of design variables like instructional strategy, process of communication, collaboration, presence of an embodied agent, etc; in the assessment and interpretation of emotional experience. A data collection exercise was undertaken in which eight participants were video-recorded while doing two computer-based learning tasks. About four hours of data was collected which underwent three annotation levels to finally get samples of the six emotion groups. The pre-selected emotion categories were validated during the annotation process except for the addition of *surprise* which did not feature in the original list of relevant emotions. *Surprise* was added to the list of domain relevant affect states because of its frequent occurrence in the data as noted by the coders. The set of affect states thus represents the range of emotions observed in the collected video data. Furthermore, the proportion of labelled instances showed the predominance of confusion followed by surprised, interested, happy, bored and annoyed. A detailed description of the data collection and annotation process appears elsewhere in (Afzal & Robinson, 2009). Note that the emotions groups of annoyed and anxious had very few representative samples to merit proper statistical analyses and were therefore not included in the subsequent analysis. The compiled dataset was used as the ground-truth for the training of a fully automatic parallel inference system designed to continuously and unobtrusively model emotions in real-time, as described in the following sections.

Representation and Measurement of Facial Motion

Machine perception of affect can be posed as a pattern recognition problem, typically classification or categorisation, where the classes or categories correspond to the different emotion groups. Determining an optimal feature representation is then crucial to overall classifier design. Defining features implies developing a representation of the input pattern that can facilitate classification. Domain knowledge and human instinct play an important role in identifying such descriptors. Although a large body of work dealing with human perception of facial expressions exists, there have been very few attempts to develop objective methods for quantifying facial movements (Essa, 1997). One of the most significant works in this area is that of Ekman & Friesen (1978) who have devised a system for objectively coding all visually distinguishable facial movements called the *Facial Action Coding System* (FACS).

FACS associates facial expression changes with the actions of the muscles that produce them and by enumerating 44 action units (AUs) it encodes all anatomically possible facial expressions, singly or in combination. Since the AUs are purely descriptive measures of facial expression changes, they are independent of interpretation and provide a useful grammar for use as feature descriptors in expression studies as this. FACS remains a popular method for measuring facial behaviour and continues to have normative significance in automatic facial expression analysis as the only psychometrically rigorous and comprehensive grammar of facial actions available (Cohn, 2006).

The 2D face model (see Figure 1) of the Nevenvision FaceTracker is used to characterize the facial motion in terms of AUs. This FaceTracker is a state-of-art facial feature point tracking technology and requires no manual pre-processing or calibration. It is resilient to limited out-of-plane motion, can deal with a wide range of physiognomies

and can also track faces with glasses or facial hair. The FaceTracker uses a generic face template to capture the movement of 22 facial feature points over the video sequences. The displacement of these feature points over successive frames encodes the motion pattern of the face AUs in a feature vector. To remove the effects of variation in face scale and projection, the distance measurements are normalized with respect to a positional line connecting the inner eyes in the first frame. Statistically, the representative values of AUs in terms of local concentration (median) and dispersion (standard deviation) are selected as parameters, along with the first temporal derivative corresponding to speed as an additional attribute. The inclusion of speed helps qualify the dynamic information in expression changes and is found to increase the interpretive power and performance of classifiers (Tong, Liao, & Ji, 2007; Pantic & Patras, 2006; Ambadar, Schooler, & Cohn, 2005).

Preliminary statistical analysis using WEKA followed by a comparison of two popular class binarisation strategies namely, the one-versus-all approach (OvA), and the pairwise or round robin approach (AvA), indicated enhanced classification performance using OvA. *Class binarisation* reduces the complexity of multi-class discrimination by transforming the original multi-class learning problem $\gamma=\{1,2,\dots,k\}$ into a series of binary problems and evaluates the overall performance by combining the multiple outputs (Littlewort, Bartlett, Fasel, Susskind, & Movellan, 2006). OvA is the most common binary classification approach based on the assumption that there exists a single (simple) separator between a class and all others. Learning proceeds by learning k independent binary classifiers, one for each class, where the positive training examples are those belonging to the class while the negative examples are formed by the union of all other classes (Park & Furnkranz, 2007; Har-Peled, Roth, & Zimak, 2003). OvA classifiers operate by a winner-takes-all strategy so that a new example is assigned to the class corresponding to the maximum output value from the k binary classifiers. The OvA scheme is powerful because of its conceptual simplicity and comparative performance relative to other binarisation methods but at lower computational costs (Rifkin & Klautau, 2004). Applying OvA strategy therefore creates six binary classifiers, each differentiating a class from all others. Positive and negative samples of relevant emotion classes are randomly sub-sampled to learn each binary classifier. From an application perspective, a classifier should also be able to deal with real-time data input and be able to model the temporal evolution of facial expressions. To address this, we now describe the classification system that uses a class of dynamic probabilistic network to model the temporal signatures of the six emotion classes under study using an OvA design.

Discriminative HMMs

Hidden Markov Models (HMMs) are a popular statistical tool for modeling and recognition of sequential data and have been successfully used in applications like speech recognition, handwriting recognition, gesture recognition and even automatic facial expression classification (Rabiner, 1989). Based on whether the observations being modelled are discrete or continuous, HMMs can be constructed as having discrete or continuous output probability densities. Since it is intuitively more advantageous to use *continuous HMMs* (CHMM) to model continuous observations, we use CHMMs to model the temporal patterns of the emotion classes under study. Following OvA design, we use HMMs in a discriminatory manner which implies learning one HMM per class, running all HMMs in parallel and choosing the model with the highest likelihood as the most likely classification for a sequence. This way an HMM models the temporal signature of each emotion class so that the likelihood that an unseen sequence is emitted by each of the models can be estimated and be classified as belonging to the model most likely to have produced it (Oliver & Horvitz, 2005; Cohen, Sebe, Garg, Chen, & Huang, 2003).

Thus, a bank of HMMs is learned using the Baum-Welch algorithm (Rabiner, 1989) over the sample sequences. During training, the Gaussian mixtures with diagonal covariance are used and the initial estimates of state means and covariance matrices are found by k-means clustering. For classification, all HMMs are run in parallel and the forward-backward procedure (Rabiner, 1989) is used to select the model with the highest likelihood as the true class. See Figure 1 for illustration. The observation vector for the HMMs consists of the position and speed parameters sampled over a sliding-window of five frames. This results in a multi-dimensional feature vector characterizing a filtered pattern sequence of the temporally evolving facial and head motions. PCA is used to extract salient features and reduce dimensionality.

The overall classification accuracy is estimated by averaging the true positive rate using tenfold cross-validation. To determine the best performance empirically, recognition accuracies are computed by varying the free parameters - the number of states and the number of Gaussian mixtures. Table 4 shows the detailed confusion matrix for the best

classification achieved. Overall, for a mean false positive rate of just 1.01% the best average accuracy of 94.96% is obtained with eleven states and four Gaussian mixtures. Happy and surprised attain perfect true positive rates while others show satisfactory recognition. Individual classes attain optimal performance at varying number of states and mixtures suggesting that individual emotions have their own temporal signatures and can be modeled by aligning them along their optimal topologies. This, along with an assessment of the generalization ability, needs to be determined in future work as it requires evaluation of the system on a database that is comparable at least in terms of context and recording conditions.

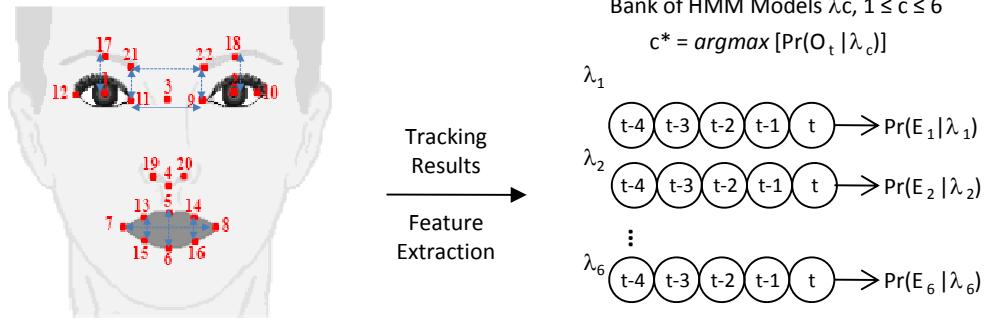


Figure 1. Feature point measurements fed to the bank of discriminative HMMs

Table 4. Best performance of discriminative HMMs.

| Predicted → | | bored | confused | happy | interested | neutral | surprised | total | TP % |
|-------------|--|-------|----------|-------|------------|---------|-----------|-------|-------|
| ↓ Actual | | 15 | 1 | 0 | 0 | 0 | 0 | 16 | 93.8 |
| bored | | 0 | 57 | 1 | 0 | 0 | 1 | 59 | 96.6 |
| confused | | 0 | 0 | 32 | 0 | 0 | 0 | 32 | 100.0 |
| happy | | 0 | 4 | 0 | 27 | 0 | 0 | 31 | 87.1 |
| interested | | 0 | 0 | 1 | 0 | 24 | 1 | 26 | 92.3 |
| neutral | | 0 | 0 | 0 | 0 | 0 | 32 | 32 | 100.0 |
| surprised | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| total | | 15 | 62 | 34 | 27 | 24 | 34 | 196 | 95.0 |
| FP % | | 0.0 | 3.6 | 1.2 | 0.0 | 0.0 | 1.2 | 1.0 | |

Summary and Conclusions

A consistent theme that emerges from education literature is that teaching and learning are essentially emotional practices. Learners experience a wide range of both positive and negative emotions, and these influence their cognitive functioning and performance. Access to emotions is then important to ensure optimal learning, more so in the case of computer-based learning environments where the learner's motivation is an important determinant of engagement and success. However, automatic measurement of affect is a challenging task. Emotions consist of multiple components that may include intentions, action tendencies, appraisal, other cognitions, central and peripheral changes in physiology, and subjective feelings. As a result they are not directly observable and can only be inferred from expressive behaviour, self-report, physiological indicators, and context (Cohn, 2006).

This paper has outlined the problem space with respect to the application of affect-sensitive technologies in computer-based learning. Building on a discussion of studies highlighting the relevance of emotions in learning, the different techniques for measuring emotions and recent advances in automatic recognition and/or prediction of affect in learning contexts were discussed. Six categories of pertinent affect states were identified; the visual modality for affect modelling was preferred given the requirements of a viable measurement technique; and a bottom-up analysis approach based on context-relevant data was adopted. Finally, a dynamic classification system using a bank of discriminative HMMs was described while the underlying differences in the temporal signatures of the individual affect states was also highlighted. Trained on the compiled corpus, it is designed to model multiple emotions simultaneously in real-time using automatic facial feature point tracking and will be optimized in future work on dataset(s) from potential learning contexts.

Acknowledgements

This research was supported by the Gates Cambridge Scholarships and the Overseas Research Studentship of the University of Cambridge.

References

- Afzal, S., & Robinson, P. (2009). Natural Affect Data - Collection and Annotation in a Learning Context. *Affective Computing & Intelligent Interaction*. Amsterdam.
- Alexander, S. T., Hill, S., & Sarrafzadeh, A. (2005). How do Human Tutors Adapt to Affective State? *Proceedings of User Modelling*. Edinburgh, Scotland.
- Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, 16 (5), 403-410.
- Amershi, S., Conati, C., & Maclare, H. (2006). Using Feature Selection and Unsupervised Clustering to Identify Affective Expressions in Educational Games. Workshop on Motivational and Affective Issues in ITS. *Intelligent Tutoring Systems*, (pp. 21-28).
- Baker, R. S. (2007). Modeling and Understanding Students' Off-Task Behaviour in Intelligent Tutoring Systems. *CHI* (pp. 1059-1068). San Jose, USA: ACM.
- Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. (2004). *Mind Reading: The Interactive Guide to Emotions*. London: Jessica Kingsley Publishers.
- Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13 (6), 4-16.
- Boekaerts, M. (1992). The adaptable learning process: Initiating and maintaining behavioural change. *Journal of Applied Psychology: An International Review*, 41, 377-397.
- Boekaerts, M. (2002). The online motivation questionnaire: A self-report instrument to assess students' context sensitivity. *New Directions in Measures and Methods*, 77-120.
- Boekaerts, M. (2003). Towards a model that integrates motivation, affect and learning. *British Journal of Educational Psychology Monograph. BJEP Monograph Series II, Number 2 - Development and Motivation*, 1 (1), 173-189.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience and School*. Washington, DC: National Academy Press.
- Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). Facial Expression Recognition from Video Sequences: Temporal and Static Modelling. *Computer Vision and Image Understanding*, 91, 160-187.
- Cohn, J. F. (2006). Foundations of Human Computing: Facial Expression and Emotion. *International Conference on Multimodal Interfaces (ICMI)*. Banff, Canada: ACM.
- Conati, C. (2002). Probabilistic Assessment of User's Emotions in Educational Games. *Journal of Applied Artificial Intelligence*, 16, 555-575.
- Conati, C., & Maclare, H. (2004). Evaluating a Probabilistic Model of Student Affect. *7th International Conference on Intelligent Tutoring Systems (ITS)*. Maceio, Brazil.
- Conati, C., & Zhou, X. (2002). Modelling Students' Emotions from Cognitive Appraisal in Educational Games. *Paper presented at the International Conference on Intelligent Tutoring Systems*, June 2-7, Biarritz, France and San Sebastian, Spain.
- Cowie, R. (2009). Perceiving emotion: towards a realistic understanding of the task. *Philosophical Transactions of the Royal Society B: Biological Science*, 364 (1535), 3515.
- Cowie, R., & Cornelius, R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, 40, 5-32.
- Cowie, R., & McKeown, G. (2009). The challenges of dealing with distributed signs of emotion: theory and empirical evidence. *Affective Computing and Intelligent Interaction (ACII)* (pp. 351-356). Amsterdam: IEEE.
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine*, 18 (1), 32-80.
- Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*, 29, 241-250.

- de Vicente, A., & Pain, H. (2002). Informing the Detection of the Students' Motivational State: An Empirical Study. *Paper presented at the International Conference on Intelligent Tutoring Systems*, June 2-7, Biarritz, France and San Sebastian, Spain.
- de Vicente, A., & Pain, H. (1998). Motivation Diagnosis in Intelligent Tutoring Systems. In B. P. Goettl, C. Halff, C. L. Redfield, & V. J. Shute (Ed.), *Intelligent Tutoring Systems*, (pp. 86-95). Texas.
- Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., McRorie, M., et al. (2007). The HUMAINE Database: Addressing the Collection and Annotation of Naturalistic and Induced Emotional Data. *Affective Computing and Intelligent Interaction* (pp. 488-500). Springer.
- D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. W., & Graesser, A. C. (2005). Integrating Affect Sensors in an Intelligent Tutoring System. Retrieved May 1, 2011 from <http://affect.media.mit.edu/pdfs/05.dmello-et.al.pdf>.
- D'Mello, S., & Graesser, A. (2007). Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments. *International Conference on Artificial Intelligence in Education*. Los Angeles.
- D'Mello, S., Picard, R. W., & Graesser, A. (2007). Towards An Affect-Sensitive Auto-Tutor. *IEEE Intelligent Systems*, 22 (4), 53.
- D'Mello, S., Taylor, R., Davidson, K., & Graesser, A. (2008). Self Versus Teacher Judgements of Learner Emotions During a Tutoring Session with AutoTutor. *Intelligent Tutoring Systems (ITS)*, (pp. 9-18).
- du Boulay, B., & Luckin, R. (2001). Modelling Human Teaching Tactics and Strategies for Tutoring Systems. *International Journal of Artificial Intelligence in Education*, 12, 235-256.
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Essa, I. A. (1997). Coding, Analysis, Interpretation, and Recognition of Facial Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (7), 757-763.
- Har-Peled, S., Roth, D., & Zimak, D. (2003). Constraint Classification for Multiclass Classification and Ranking. (S. Becker, & K. Obermayer, Eds.) *Advances in Neural Information Processing Systems*, 15.
- Heylen, D., Ghijssen, M., Nijholt, A., & Akker, R. (2005). Facial Signs of Affect During Tutoring Sessions. In J. Tao J, & R. W. Picard (Ed.), *Affective Computing and Intelligent Interaction. Lecture Notes in Computer Science*, 3784, 24-31.
- Hudlicka, E. (2003). To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Studies*, 59, 1-32.
- Jaques, P. A., & Vicari, R. M. (2007). A BDI approach to infer students' emotions in an intelligent learning environment. *Computers & Education*, 49, 360-384.
- Jarvenoja, H., & Jarvela, S. (2005). How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach. *Learning and Instruction*, 15 (5), 465-480.
- Immordino-Yang, M. H., & Damasio, A. (2007). We Feel, Therefore We Learn: The Relevance of Affective and Social Neuroscience to Education, Mind, Brain and Education. *Mind, Brain and Education*, 1 (1), 3-10.
- Kapoor, A., Burleson, W., & Picard, R. W. (2007). Automatic Prediction of Frustration. *Journal of Human-Computer Studies*, 65 (8), 724-736.
- Kapoor, A., & Picard, R. W. (2005). Multimodal Affect Recognition in Learning Environments. *13th Annual ACM International Conference on Multimedia*. Singapore.
- Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. *IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges*, (pp. 43-48). Madison.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. (S. P. Lajoie, & S. J. Derry, Eds.) *Computers as Cognitive Tools*, 75-105.
- Liao W, W., Zhang, W., Zhu, Z., Ji, Q., & Gray, W. D. (2006). Toward a decision-theoretic framework for affect recognition and user assistance. *International Journal of Human-Computer Studies*, 64, 847-873.
- Lisetti, C., & Schiano, D. (2000). Facial expression recognition: Where Human Computer Interaction, Artificial Intelligence and Cognitive Science Intersect. *Pragmatics and Cognition*, 8 (1), 185-235.
- Litman, D., & Forbes, K. (2003). Recognising emotions from student speech in tutoring dialogues. *Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, New York: IEEE, 25-30.
- Littlewort, G., Bartlett, M. S., Fasel, I., Susskind, J., & Movellan, J. (2006). Dynamics of Facial Expression Extracted Automatically from Video. *Image and Vision Computing*, 24 (6), 615.
- Mavrikis, M., Maciocia, A., & Lee, J. (2007). Towards Predictive Modelling of Student Affect from Web-Based Interactions. *Proceedings of the International Conference on Artificial Intelligence in Education*. Amsterdam: IOS Press, 169-176.

- Merill, D. C., Reiser, B. J., Trafton, J. G., & Ranney, M. (1992). Effective Tutoring Techniques: A Comparison of Human Tutors and Intelligent Tutoring Systems. *Journal of the Learning Sciences*, 2, 277-305.
- Meyer, D. K., & Turner, J. C. (2002). Discovering emotion in classroom motivation research. *Educational Psychologist*, 37, 107-114.
- O'Regan, K. (2003). Emotion and e-learning. *Journal of Asynchronous Learning Networks*, 7 (3), 78-92.
- Oliver, N., & Horvitz, E. (2005). A Comparison of HMMs and Dynamic Bayesian Networks for Recognizing Office Activities. *Lecture Notes in Computer Science*, 3538, 199-208.
- Ortony, A., Clore, G. L., & Collins, A. (1998). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
- Pantic, M., & Patras, I. (2006). Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences. *IEEE Transactions on Systems, Man, and Cybernetics*, 36 (2), 433-449.
- Pantic, M., & Rothkrantz, L. J. (2003). Toward an Affect-Sensitive Multimodal Human-Computer Interaction. *Proceedings of the IEEE*, 91 (9), 1370-1390.
- Park, S.H., & Furnkranz, J. (2007). Efficient Pairwise Classification. *Lecture Notes in Computer Science*, 4701, 658-665.
- Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction*, 15, 497-506.
- Pekrun, R. (1992). The Impact of Emotions on Learning and Achievement: Towards a Theory of Cognitive/Motivational Mediators. *Applied Psychology*, 41, 359-376.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91-105.
- Peter, C., & Herbon, A. (2006). Emotion representation and physiology assignments in digital systems. *Interacting with Computers*, 18, 139-170.
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., et al. (2004). Affective Learning-a manifesto. *BT Technology Journal*, 22, 253-269.
- Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77 (2), 257-285.
- Rifkin, R., & Klautau, A. (2004). In Defense of One-Vs-All Classification. *Journal of Machine Learning Research*, 5, 101-141.
- Sarrafzadeh, A., Fan, C., Dadgostar, F., Alexander, S., & Messom, C. (2004). Frown gives game away: Affect sensitive tutoring systems for elementary mathematics. *IEEE Conference on Systems, Man and Cybernetics*. The Hague.
- Scherer, K. L. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44 (4), 695-729.
- Schutz, P. A., Hong, J. Y., Cross, D. I., & Obson, J. N. (2006). Reflections on Investigating Emotion in Educational Activity Settings. *Educational Psychology Review*, 18, 343-360.
- Tong, Y., Liao, W., & Ji, Q. (2007). Facial action unit recognition by exploiting their dynamic and semantic relationships. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (10), 1683-1699.
- Turner, J. E., Husman, J., & Schallert, D. L. (2002). The importance of students'-goals in their emotional experience of academic failure: Investigating the precursors and the consequences of shame. *Educational Psychologist*, 92, 548-573.
- van Vuuren, S. (2006). Technologies that power pedagogical agents and visions for the future. Retrieved May 1, 2011 from http://www.bltek.com/images/workshops/2004/ed_tech_pedagogical_agents_022206.pdf.
- Whitehall, J., Bartlett, M., & Movellan, J. (2008). Automatic Facial Expression Recognition for Intelligent Tutoring Systems. *Paper presented at the IEEE Computer Vision and Pattern Recognition Conference*, June 23-28, Anchorage, Alaska, USA.
- Wosnitza, M., & Volet, S. (2005). Origin, direction and impact of emotions in social online learning. *Learning and Instruction*, 15 (5), 440-464.
- Zakharov, K., Mitrovic, A., & Johnston, L. (2008). Towards Emotionally-Intelligent Pedagogical Agents. *Lecture Notes in Computer Science*, 5091, 19-28.
- Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 31 (1), 39-58.

A Learning Content Authoring Approach based on Semantic Technologies and Social Networking: an Empirical Study

Saša Nešić¹, Dragan Gašević², Mehdi Jazayeri³ and Monica Landoni³

¹Dalle Molle Institute for Artificial Intelligence, Lugano, Switzerland // ²School of Computing and Information Systems, Athabasca University, Canada // ³Faculty of Informatics, University of Lugano, Lugano, Switzerland // sasa@idsia.ch // dgasevic@acm.org // mehdi.jazayeri@usi.ch // monica.landoni@usi.ch

ABSTRACT

Semantic web technologies have been applied to many aspects of learning content authoring including semantic annotation, semantic search, dynamic assembly, and personalization of learning content. At the same time, social networking services have started to play an important role in the authoring process by supporting authors' collaborative activities. Whether semantic web technologies and social networking improved the authoring process and to what extent they make authors' life easier, however, remains an open question that we try to address in this paper. We report on the results of an empirical study based on the experiments that we conducted with the prototype of a novel document architecture called SDArch. Semantic web technologies and social networking are two pillars of SDArch, thus potential benefits of SDArch naturally extend to them. Results of the study show that the utilization of SDArch in authoring improves user' performances compared to the authoring with conventional tools. In addition, the users' satisfaction collected from their subjective feedback was also highly positive.

Keywords

Empirical study, Learning content authoring, Semantic web technologies, Social networking

Introduction

Authoring of learning content completely from scratch has always been a difficult and time-consuming task. Current research has shown that most authors reuse and modify existing learning content, available in their own archives or on the Web (Betty & Allard, 2004) rather than authoring new content from scratch. Therefore, if the main goal of learning content is for teaching and learning, the second goal should be its reuse. The reuse process requires a meaningful way to search and retrieve the appropriate content. Extensive research has been carried out lately to enhance the reusability of learning content by leveraging the semantic web technologies for standardization and semantic annotation of learning content components (Duval et al., 2001; Jovanović et al., 2006). While these efforts have demonstrated some significant potential to improve the current state of the authoring of learning content, there are still some important issues to be addressed. Firstly, ontology-based semantic annotation approaches (Uren et al., 2006) represent a step ahead comparing to the standardized metadata annotation, but the full potential of the semantic search will be achieved when learning content components can be efficiently searched by means of semantic annotations as well as structural and semantic relationships between them. Thus, not only the semantic annotation, but also the framework for linking learning content components and adding logical assertions over linked components is necessary. Secondly, most of the existing learning content is isolated in huge, centralized repositories with restricted access, which is opposite of trends of the emerging Web 2.0 (Berners-Lee et al., 2006). Thirdly, in spite of a number of different learning object (LO) models and LO repositories (e.g., MERLOT) built on top of them as well as federated protocols (e.g., ECL and SQUI) for networks of LO repositories (e.g., GLOBE), most authors still consider conventional documents (e.g., PDFs, Word and PowerPoint) as a primary source of learning content. The main issue with conventional documents with respect to the learning content authoring is that only entire documents can be considered as resources that can be uniquely identified, searched and retrieved. In practice, however, authors usually need only document parts that are related to a certain concept and play a certain pedagogical role (e.g., illustration, definition and example) (Jovanović et al., 2006). Common selective reuse of document content is a cumbersome task requiring copy-and-paste, which is a laborious and error prone process. Finally, despite the fact that some authoring tools can provide some collaborative activities, most of conventional authoring tools are designed primarily for individual users and pay little attention to the users' social activities. Social relations between content authors and the way different authors use and interpret the same learning content could be useful information in the content authoring.

The novel semantic document architecture (such as SDArch as explained in the section "Semantic Documents") along with the underlying document representation model called a semantic document model represents our solution

to the above-discussed issues of the learning content authoring. In this paper, our main focus is on the empirical evaluation of the proposed architecture that we conducted by using the architecture prototype that we developed. Having in mind that semantic web technologies and social networking are two pillars of the new architecture, we can consider the conducted evaluation as the evaluation of the use of these two types of technologies in the learning content authoring. So far, the use of semantic web technologies in learning content authoring has been reported in many studies (Duval et al., 2001; Dodero et al., 2005; Jovanović et al., 2006; Henze, 2005), but none of them provided any experimental data that would justify the benefits of using these technologies. In other words, it is still unclear up to what extent the use of these technologies can improve the authors' effectiveness, efficiency and satisfaction compared to conventional learning content authoring approaches. The evaluation that we conducted was designed to compare the authors' *effectiveness*, *efficiency* and *satisfaction* in using the SDArch prototype for authoring of course material compared to conventional authoring tools. Results of quantitative and qualitative measures that we applied in the evaluation showed promising improvements that the use of semantic technologies and social networking brings to the learning content authoring.

The paper is organized as follows. In Section 2, we first describe a motivational authoring scenario that relies on the use of services that are enabled by semantic web and social networking technologies. Subsequently, in Section 3, we introduce the notion of semantic documents and discuss both the semantic document model and the design of the proposed semantic document architecture. Section 4 provides details of the SDArch tools and services that are essential for the realization of the given motivational authoring scenario. In Section 5, we first discuss the design of our evaluation study, then explain how the evaluation study was conducted, and eventually present and discuss obtained evaluation results. Section 7 outlines relevant related works, and Section 8 concludes the paper.

Authoring of Course Material – Motivational Scenario

Let us suppose that Mark is a university professor who teaches ‘Software Architecture and Design’ course. For each topic in the course Mark usually prepares presentation slides that he uses during his class. The next topic to be presented in the course is ‘Software Design Patterns’. Mark has the presentation on this topic from previous year, but he does not want to reuse it as it is. In order to prepare as good presentation as possible, with up to date information, Mark plans to consider the existing presentation, then presentations on the same topic used by his colleagues at other universities, and some other articles related to the topic from his archive as well as those of his colleagues. As usual, Mark is going to use PowerPoint to prepare the presentation, as he is most confident and familiar with it. However, this time his PowerPoint is extended with a set of tools that provide him a range of new, novel services, which we could categorize into four groups.

The first group contains the social networking services that allow Mark and his colleagues who teach the same topic to organize themselves into an online social networking group dedicated to that topic. For all the member the group manages their subjective, self-assessed expertise on the topic as well as objective, quantitative data that shows the members' expertise such as a number of their citations in the topic's related literature and their ratings within the group formed based on the votes of the other members. Moreover, these services provide functionalities for managing Mark's profile and allow him to specify his preferences regarding the automatic selection of document content for reuse. The examples of these preferences include an ordered list of preferred network members, an ordered list of preferred document formats, and information if the user prefers content that is often reused, recently modified content or content with many versions.

The second group contains services that enable Mark to transform his office documents (i.e., Word and PowerPoint) into a novel document representation form that is completely open and queryable and that encapsulates document content into reusable, uniquely identified and semantically annotated data units. Moreover, these services enable Mark to store transformed documents either on his laptop or to publish them onto a shared document repository of the social networking group.

The third group contains services that enable Mark to search local and shared documents for document units not only based on their content but also their semantics (i.e., semantic search). Moreover, these services take into account Mark's preferences stored in his user profile and recommend to him those search results that correspond well with the preferences. Before reusing some document units, Mark can preview their content and browse available annotations. Once he decides which document unit to reuse, he can fetch the document unit automatically into a new

document without a need to obtain a whole document that the document unit originates from. In addition, these services observe Mark's behavior and track the data of his interaction with document units (e.g., times when he browses and reuses document units) and the way he modifies reused document units to fit to a new context.

The fourth and the last group of services provide Mark the ability to navigate across collections of documents stored on his laptop as well as those documents from the social network repository by following explicit semantic links between semantically related document units. The explicit semantic links are enabled by the new document representation form and are established based on the conceptualized semantics of document units.

To summarize, the novel document representation and the envisioned services that will run on top of it will enable Mark and his colleagues:

- to form a social network around a given topic of interest;
- to transform their local documents in a new form that will enable semantic integration (i.e., semantic annotation and linking) of related data kept in different documents;
- to share such transformed documents within the social network, and thus, semantically integrate related document data that originate from different users;
- to semantically search local and shared collections of the semantically integrated documents for desired data; and
- to navigate across local and shared document collections by following semantic links between document units and thus discover more data units of their interest.

Semantic Documents

In order to bring desktop documents closer to the motivational authoring scenario, we introduced a new form of documents, namely semantic documents, described by *a semantic document model (SDM)*, and designed a supporting software architecture called *Semantic Document Architecture (SDArch)*. Semantic documents enable unique identification, semantic annotation, and semantic linking of fine-grained data units regardless of whether they belong to the same or different documents. Moreover, semantic documents enable semantic links to be established between semantically related data units stored on personal desktops and published on the Web into shared repositories of online social network communities. Therefore, semantic documents have potential to integrate data from desktop documents into a unified desktop information space. At the same, semantic documents can fill the gap between the desktop information space and the information space of the online social network communities. Novice processes such as semantic document search and navigation, which will run on such integrated information space, will improve the effectiveness and efficiency of desktop users in carrying out their daily tasks.

In the rest of this section, we first outline the main features of SDM, and then, describe the SDArch design. In the next section, we take a closer look at SDArch tools and underlying services that are essential for the given motivational authoring scenario.

Semantic Document Model – SDM

SDM defines semantic documents as composite information resources composed of uniquely identified, semantically annotated, and semantically interlinked document units (DUs) of different granularity (Nešić, 2009). Each semantic document is characterized by unique, permanent machine-processable (MP) representation and a number of temporal human-readable (HR) representations rendered from the MP representation. The formal specification of SDM is done by the SDM ontology, which consists of four parts: the core part, the annotation part, the semantic-linking part and the change-tracking part.

The core part of the SDM ontology provides a vocabulary (classes and properties), which defines possible types of DUs and structural relationships among them. The two main DU types are atomic DUs and composite DUs. An atomic DU contains a single unit of raw digital content that exists as a physical entity independently of the document unit it belongs to and cannot be disaggregated into smaller units. A composite DU aggregates a number of atomic or other composite DUs and organizes them in a given order.

The annotation part of the SDM ontology provides the annotation vocabulary that describes possible types of the DUs' annotations as well as provides the annotation interface (i.e., properties) for linking annotations to DUs. The annotation interface is designed, so that all DUs' annotations, regardless of the annotation type, are linked to DUs in the same way. The current version of the annotation vocabulary contains concepts and properties that specify the three types of DU annotations: *semantic annotation*, *social-context annotation* and *pedagogical annotation*. The semantic annotations refer to concepts from domain ontologies that represent the conceptualization of the information/knowledge held by DUs. The social-context annotations (Nešić et al., SoSEA 2009) capture relevant information about the user actions such as browsing, reusing and modification that are performed to DUs in a given social context. Finally, if semantic documents hold some learning content, then their DUs can be annotated by the pedagogical annotations that we introduced to model potential pedagogical roles (e.g., *abstract*, *introduction*, *conclusion*, *definition*, *explanation*, *description*, *illustration*, *example* and *exercise*) of the DUs.

The semantic-linking part of the SDM ontology defines the interface for linking semantically related DUs. Semantic links are determined by the ontological concepts that conceptualize shared semantics between the linked DUs. The semantic links enable the semantic navigation across integrated collections of semantic documents and thus help in the discovery of semantically related DUs. Together with the semantic annotations, the semantic links constitute the semantic layer of semantic documents.

The change-tracking part of the SDM ontology provides a vocabulary that defines possible changes of DUs as well as changes of the whole semantic document.

Semantic documents, that is, instances of SDM, employ HTTP-dereferenceable URIs to identify DUs and the Resource Description Framework (RDF) data model to represent structural and semantic links between them. The use of the HTTP-dereferenceable URIs and the RDF data model is inline with the Linked Data principles, so that semantic documents can be seamlessly integrated to the Linked Open Data cloud (Berners-Lee et al., 2006) and further to the envisioned Semantic Web. Moreover, the conceptualization of the document semantics by ontological concepts and the establishment of explicit semantic links between related DUs, can lead to the creation of a sufficient amount of semantically integrated data. This creation is necessary for the Semantic Web to succeed.

Semantic Document Architecture - SDArch

In order to support semantic document management and to enable users to take advantage of semantic documents, we designed the supporting software architecture called the semantic document architecture or SDArch. SDArch is a three-tier, service-oriented architecture (see Figure 1) composed of the data layer, the service-oriented middleware, and the user interface layer.

The data layer contains the semantic document repository that is composed of the RDF and the binary data repositories, and equipped with the concept and text indexes. The RDF repository stores RDF instances of semantic documents. A binary content of semantic document units is kept separately from RDF document representations and stored in the binary data repository. SDArch maintains the single concept index that enables the semantic document search over RDF data and the single text index that enables the full-text search over document binary data. Both indexes are updated every time a new document is added to or removed from the repository. In addition, the repository exposes remotely accessible SPARQL endpoint, so that SPARQL queries can also be sent from remote machines over HTTP.

The service-oriented middleware provides the service registry and establishes the communication protocol among the SDArch services and between the SDArch services and the user interface. In the actual design, the SDArch (Nešić et al., SEKE 2010) functionalities are encapsulated into five services: 1) the semantic document authoring, 2) the semantic document search and navigation, 3) the user profile management, 4) the social network management, and 5) the ontology management services. Among other functionalities, the SDArch services provide most of the functionalities intended for the realization of the motivational authoring scenario. Potential new functionalities can be added to SDArch by registering new services into the SDArch middleware.

The presentation layer is the top layer of SDArch, which provides the user interface for the SDArch services. According to the service-oriented nature of SDArch, the presentation layer is technology- and platform-independent.

It can contain web-based applications, desktop-based applications and mobile phone applications. In the prototype that we have developed, we focused on extending the existing document-authoring suites, instead of creating completely new tools. In this way, we enable users to take advantage of the SDArch services, while still working within familiar environments. As an example, we extended MS Office with a set of tools that we named 'SemanticDoc' tools. We chose MS Office, mostly because of its wide usage and popularity.

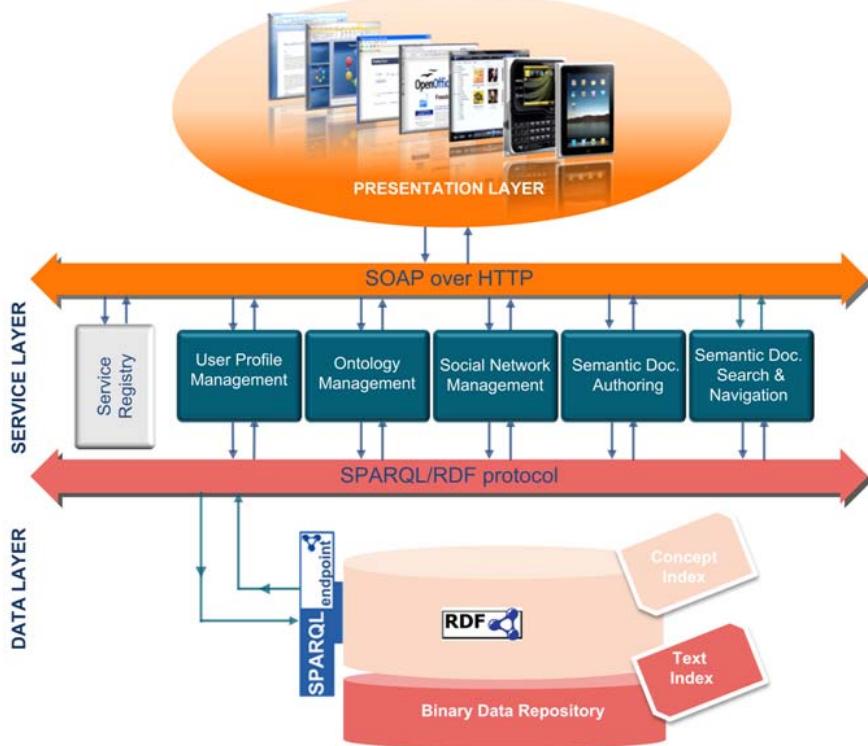


Figure 1. Illustration of the SDArch architecture

SemanticDoc Tools

SemanticDoc tools enable Microsoft (MS) Office users to take benefits of the SDArch services directly from MS Office (i.e., MS Word and MS PowerPoint). In other words, they provide access to the SDArch services from within MS Office. Since SDArch enables users to share their semantic documents and to form a social network around shared documents, the SemanticDoc tools actually turn MS Office into a social environment.

The tools are grouped and accessible through several toolboxes. Each toolbox contains a set of tools that provide the interface for interacting with a certain group of SDArch services. In this paper, our focus is on the social network manager, the document recommender and the semantic document browser tools, as they are essential for the motivational authoring scenario that we want to evaluate. More information, snapshots and demos of all SemanticDoc tools can be found at our project Web page (www.semanticdoc.org).

Social Network Manager

The SDArch social network management service provides functionalities for organizing SDArch users into a social network and sharing their semantic documents by publishing the RDF representations of the documents into the network's shared RDF repository. Moreover, the service provides functionalities for capturing interaction between the network members and the shared semantic documents and generating corresponding social-context annotation for the shared semantic documents. SemanticDoc social network manager extends the MS Office with a user interface

that enables the office users to access the SDArch social network management service. By using this tool, the office users can join the SDArch social network, and then, organize themselves within the network into groups dedicated to particular topics of interest. Every member of the SDArch social network can initiate a new group as well as join or leave an existing group. To initiate a new group, the network member needs to specify the group's topic of interest and to provide some topic-related information (e.g., the topic's short description and the list of the topic's Web references). Figure 2 shows a snapshot of the tool displaying: a) a list of all existing groups, and b) group details of a selected group. In the current prototype implementation, there is no restriction for joining existing groups, that is, all groups are available to all members of the SDArch social network.

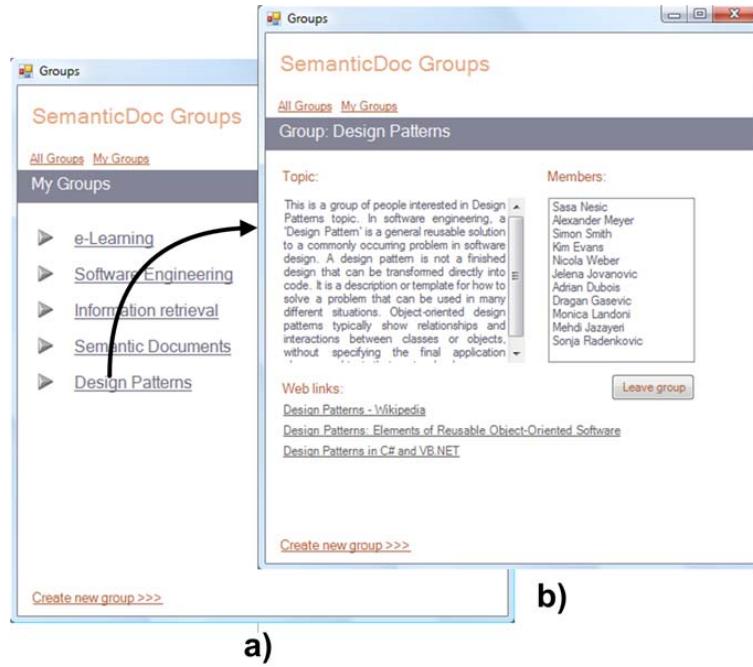


Figure 2. Social network manager: a) a list of existing groups within the SDArch social network, b) a detailed view of a selected group

Document Recommender

Document recommender tool is a starting point from where the office users start to explore semantic documents whether they are stored in a local desktop repository or in a shared repository of the SDArch social network. This tool actually provides the user interface for the personalized semantic document search (Nešić et al., SEMAPRO 2010) that is realized by the SDArch semantic document search and navigation service. The personalized semantic document search is founded on the utilization of the conceptualized DU semantics (i.e., DU semantic annotations), DU social-context annotations, and user preferences held in the SDArch user profile (Nešić et al., SoSEA 2009).

The user interface of the document recommender (see Figure 3) enables the office user to specify the following search parameters. Firstly, the user specifies which semantic document repository will be searched (i.e., local or shared). Secondly, the user specifies the user query in a form of the free-text keyword query. The tool offers the auto-completion keyword suggestion support, which helps the user while specifying the query. Suggested terms are concept labels from domain ontologies that have been used for the semantic annotation and indexing of the semantic documents from the specified repository. Thirdly, the user selects the content type of the desired document units (i.e., text, image, audio or video). Fourthly, if searching for learning content, the user specifies a pedagogical role (Jovanovic et al., 2006) of the document units (e.g., definition, example and illustration). Finally, the user specifies the search type: the semantic document search or the full-text search. If the user selects the semantic document search, the keyword query will be transformed into a semantic query and then executed by the service against the concept index of the selected repository. Otherwise, the service executes the initial keyword query against the text index of the repository. When the search is done, the document recommender displays previews of the retrieved document units to the user. Figure 3 gives snapshots of the document recommender displaying: a) the search form

and previews of the top ranked textual document units, and b) the search form and previews of top-ranked document units of image content type. For each of the retrieved document units, the user can see the detailed view including document unit content and document unit annotation data. The detailed view is shown in the semantic document browser, which is another SemanticDoc tool that we explain next.

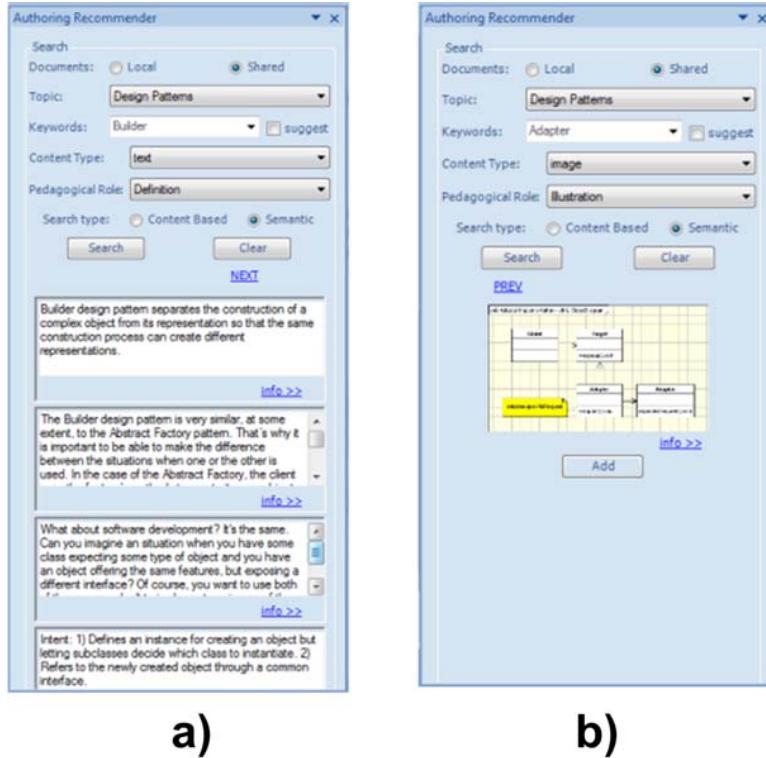


Figure 3. Document Recommender: a) an example search for textual document units, and b) an example search for document units of the image content type

Semantic Document Browser

The semantic document browser enables the user to browse details of semantic document units and to navigate across semantic documents following semantic links between the document units. In the current implementation, the browser is launched from the document recommender by clicking on previews of DUs retrieved as search results. For the next generation of the tool, we plan to enable its individual launching and the possibility to start the semantic navigation not only from the search results, but also by entering the URI of an initial document unit.

The main window of the semantic document browser (see Figure 4) is composed of two panels. The right panel displays the document unit's content, metadata (e.g., creator and creation date), and social-context annotations (e.g., the number of the document unit's reuses and the list of SDArch users who have reused the document unit). The left panel displays the document unit's semantic annotations in a form of an ordered list of ontological concepts that annotate the document unit. For each annotation concept, the user can see the concept's rank, the concept's relevance weight for the document unit, and the ontology in which the concept is defined. Moreover, for each annotation concept if there exist some document units that are linked to the document unit via semantic links determined by that concept, the browser displays the link labeled as 'browse annotated document units'. By clicking on this link, the browser invokes the SDArch semantic document search and navigation service and initiates the semantic document navigation process. The discovered document units are ordered by the strength of the semantic links between them and the initial document unit and displayed on the right panel of the browser's window.

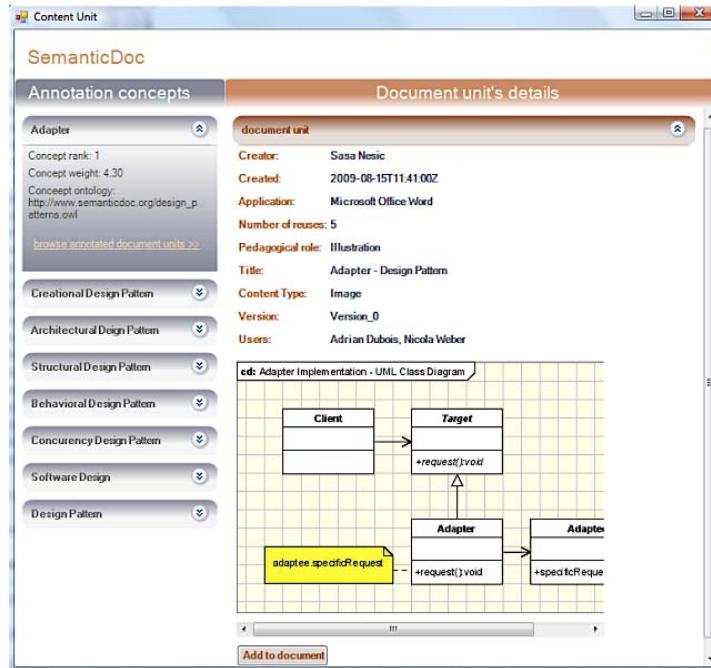


Figure 4. Semantic document browser

Empirical Evaluation of the Proposed Authoring Scenario

The goal of the empirical evaluation of the proposed authoring scenario was to investigate to which extent the authoring of course material can benefit from the proposed semantic document architecture and the underlying semantic document model. Since semantic web technologies and social networking represent two pillars of the architecture, potential benefits naturally extend to them. We formulated the evaluation hypothesis as follows:

“Using semantic web technologies and social networking results in a more **effective**, **efficient**, and **satisfactory** experience, when authoring course material compared to the conventional authoring approach.”

- With respect to user **effectiveness**, we intended to measure the accuracy and completeness with which SDArch users complete authoring tasks. In other words, how many and what tasks the users can complete successfully using the SDArch services and tools.
- With respect to user **efficiency**, we intended to measure the resources expended in relation to the accuracy and completeness with which SDArch users complete the authoring tasks. In other words, how much effort the users spend for completing these tasks using the SDArch services and tools.
- With respect to user **satisfaction**, we intended to measure the freedom from discomfort, and positive attitudes towards the use of the SDArch services and tools in authoring of course material.

Designing the Evaluation

In order to validate the evaluation hypothesis, we chose a task-based comparative evaluation (Whittaker et al., 2000) complemented with the goal-question-metrics (GQM) measurement model (Bastili et al., 1994). This implies asking test persons to perform a set of tasks in order to properly engage with two systems to be compared. In our case, one system was a conventional Windows system equipped by regular MS Office, while the other one was a Windows system featured by SDArch services and MS Office extended by the SemanticDoc tools. In the rest of the paper, we refer to these two systems as the conventional system and the SDArch system, respectively.

The set of tasks that we considered in the evaluation was composed of tasks that realize the motivational authoring scenario. It actually meant that a successful completion of the tasks should have resulted in a short PowerPoint presentation on the ‘Software Design Patterns’ topic that is composed exclusively of reused content from shared semantic documents. By considering the authors’ experience in preparing course material and in order to obtain both meaningful and feasible proof of concept we decided to keep the outcome presentation to a minimum of seven slides covering: 1) *Introduction*, 2) *Role of Design Patterns*, 3) *Design Patterns Definition*, 4) *Design Patterns Classification*, 5) *Pattern Example 1*, 6) *Pattern Example 2*, and 7) *Topic’s Conclusions*. This limited the amount of efforts required from the participants while at the same time producing an overall presentation of an appropriate quality level. All slides had to contain certain numbers of textual items and slides 5 and 6 should have also contained graphical illustrations of the chosen example patterns. Even if this sounds pretty restrictive, the aim was to set up a controlled environment where comparing effectively experiences across the new and conventional system for producing presentations, while still encouraging participants to use their creativity and expressivity.

In accordance with the evaluation hypothesis and by following the GQM model, we considered *user effectiveness*, *efficiency* and *satisfaction* as main evaluation criteria and defined qualitative and quantitative measures for them. With respect to user effectiveness, we planned to measure how effective participants were in completing the evaluation tasks. Thus, we tracked how many and which tasks participants could complete successfully by using the two systems. With respect to user efficiency, we planed to measure how efficiently participants were in completing the evaluation task. Thus, we measured the execution time, the number of mouse clicks and the number of window switches. Finally, with respect to user satisfaction, we planed to evaluate which of the compared two systems the participants liked more and why. For this purpose, we used a follow-up questionnaire that we crated by selecting a subset of questions/statements from the Perceived Usefulness and Ease of Use questionnaire (Davis, 1989) that was appropriate for our evaluation. We expected participants to implicitly and naturally refer to their previous experiences in using the conventional system when considering the performance of the SDArch system. Nonetheless, we set up a more formal comparative experiment making all participants engaged with the same documents and tasks via the two systems. This way we could extract a richer set of data to be compared in order to address our hypothesis and related research questions in terms of evaluation criteria. Table 1 summarizes the chosen evaluation methods and metrics for each of the three considered evaluation criteria.

The initial step of an empirical evaluation is the selection and recruitment of participants, whose background and abilities are representative of intended users of the system to be evaluated (Nielsen, 1993). The evaluation results will only be valid if the participants are typical users of the system, or as close to that criterion as possible. Another issue regarding the selection of the participants, which has attracted a lot of attention in the HCI community, is what should be a sufficient number of participants of the usability study. In terms of quality, Nielsen (Nielsen et al., 2003) argues that five expert users are sufficient to discover 85% of the usability problems in a system under evaluation. In our evaluation, we had six participants from three universities: University of Lugano (www.usi.ch), Switzerland; Simon Fraser University (www.sfu.ca), Canada; and University of Belgrade (www.bg.ac.rs), Serbia. All the participants were volunteers and had genuine motivation in using the new systems. Moreover, each participant had been involved in some courses covering the topic of our evaluation scenario, either as a lecturer or teaching assistant. Thus, they qualified as domain experts and final users of the system.

Table 1. Evaluation criteria and the corresponding evaluation methods and metrics

| Evaluation Criterion | Evaluation Method | Evaluation Metric |
|----------------------|--|--|
| Effectiveness | Objective – Quantitative Measure | Task Success Rates |
| Efficiency | Objective – Quantitative Measure “ “ | Task Completion Times Number of Mouse Clicks Number of Window Switches |
| Satisfaction | | 5-level Likert scale |

Conducting the Evaluation

We started the evaluation by the preparation phase whose main objectives were to create SDArch social network, to collect the evaluation document set, and to familiarize the participants with the SDArch services and tools. First, we initiated the SDArch social network by using the social network manager tool and created a *software design patterns* interest group. Then, we transformed initial 20 documents from our archive into semantic documents and added them

to the group's shared semantic document repository. After that, we invited the participants to register to the SDArch social network and to join the group. Moreover, we asked the participants to check if they had some Office (i.e., Word and PowerPoint) documents related to the software design patterns topic in their archives, and if so to transform and publish some of them to the group's semantic document repository. For the documents transformation (i.e., semantic annotation) process, all the participants were required to apply the same domain ontology, which we added to the group's ontology repository. In addition, we created a simple web-based file upload form application and asked the participants to upload the original office documents that they transformed and published to the shared repository. In that way, we also obtained the original MS Office documents, which we needed for the tasks executions with the conventional system. One week after we initiated the software design patterns group, the total number of the shared semantic documents reached 50 documents. According to our experience in preparing course material, it was a sufficient number of documents that we planned to consider in the evaluation. Therefore, we decided to organize the evaluation session and conduct the evaluation tasks.

The evaluation session consisted of two phases, namely the *observation* and *feedback* phases. In the observation phase, we were observing the participants while they were conducting the evaluation tasks and tracked their behavior by using an appropriate screen-recording software (www.techsmith.com/camtasia.asp). To avoid asking the participants to install the screen recording software and to simplify the manipulation of the recorded materials, we asked the participants to perform the evaluation tasks on our PC with remote access control. For all the participants we created accounts on the PC and enabled them to access and control it remotely. In that way the only software that the participants needed to install/enable on their computers was a remote desktop connection software. This kind of software is supported as an official feature on all new-generation, Windows operating systems (e.g., Windows XP/Vista/7), so that the participants using Windows only needed to enable it, unless they had used it before. Four out of six participants already had the software installed on their laptops and were familiar with it, while the other two did not experience and report any difficulties in using it.

The participants were split into two control groups of three participants. The first group was asked to execute the tasks first by using the conventional system and then using the SDArch system. The second group used the compared systems in the opposite order. Each participant was allowed to do the evaluation tasks within two given days at the time he preferred, but in two separate, continuous time sessions one for the conventional system and the other one for the SDArch system. The sessions started and ended by the participants activating and deactivating the screen recording software.

The observation phase was followed by the feedback phase, where we asked the participants to fill in the follow-up questionnaire. The questionnaire was composed of the following nine statements:

- S1: Using the SDArch services and SemanticDoc tools enables me to accomplish tasks more quickly;
- S2: Using the SDArch services and SemanticDoc tools increases my productivity;
- S3: Using the SDArch services and SemanticDoc tools improves the quality of the work I do;
- S4: Using the SDArch services and SemanticDoc tools makes it easier to do my work;
- S5: Overall, I find the SDArch services and SemanticDoc tools useful in my work;
- S6: Learning to operate the SDArch services and SemanticDoc tools is easy for me;
- S7: I find it easy to get the SDArch services and SemanticDoc tools to do what I want them to do;
- S8: Interaction with the SDArch services and SemanticDoc tools is clear and understandable;
- S9: Overall, I find the SDArch services and SemanticDoc tools easy to use.

The participants were supposed to rate each of the statements using 5-level Likert scale (Gediga et al., 1999), starting from 1 (strongly disagree) to 5 (strongly agree). First 5 statements from (S1-S5) were designed to gather subjective evaluation of the system usefulness. Statement S6 evaluated the system ease-of-learning. Statements S7-S9 were designed to gather subjective evaluation of the ease-of-use of the system.

Evaluation Results

By analyzing the data recorded during the observation phase, we gather indications about user effectiveness and user efficiency. With respect to user effectiveness, we tracked how many and which tasks participants completed successfully. All participants completed successfully all tasks, using both systems. In our opinion, this result is

mostly due to time-unlimited sessions and the ability of the participants to set the evaluation sessions at preferable time as well as their genuine motivation to participate in the evaluation.

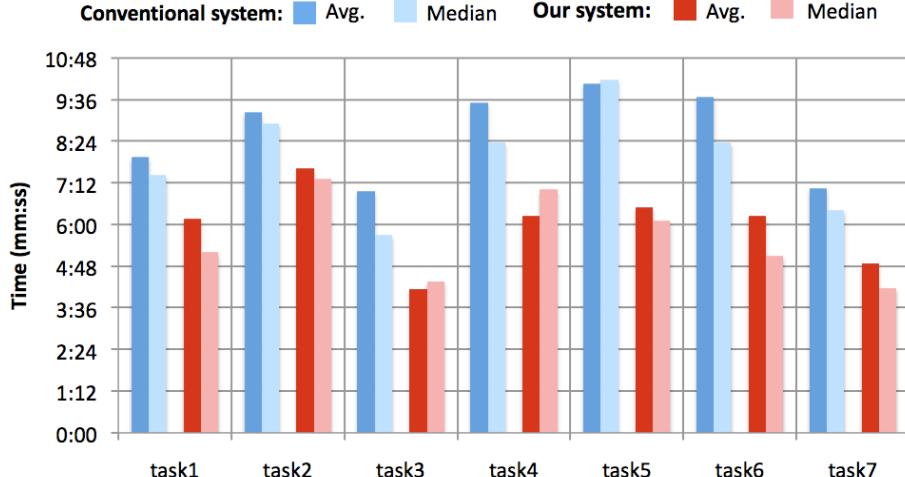


Figure 5. Average and median task completion times

With respect to user efficiency we measured the task completion times, the amount of mouse clicks and the number of window switches during the tasks executions. Table 2 and Figure 5 show the average and median task execution times of all seven considered tasks for the two compared systems. Moreover, Table 2 reports standard deviation of the task completion times for both systems, the relative performance of the participants when using the SDArch system with respect to the conventional system. For example, the relative performance of 70% indicates that the participants using the SDArch system needed 70% of the time that the participants using the conventional system needed. Finally, we performed a *t-test* (Zimmerman, 1997) to investigate on the statistical significance of the difference in the task completion times between two control groups (i.e., the participants using the SDArch system and the participants using conventional system). The results of the *t-test* (i.e., *p-values*) are shown in the last column of the table. In our case, *p-values* represent the probability that the measured task completion times for the two control groups are part of the same distribution. In general, *p-values* below 0.05 are considered statistically significant. In other words, *p-values* of 0.05 or greater indicate that there is no statistically significant difference between the results of two control groups.

Table 2. Tasks completion times' statistics

| Task | Conventional System | | | SDArch System | | | Relative Performance | | <i>t-test</i> <i>p</i> (T<=t) |
|------|---------------------|--------|------|---------------|--------|------|----------------------|--------|----------------------------------|
| | Avg. | Median | | Avg. | Median | | Avg. | Median | |
| 1 | 7:56 | 7:25 | 0:52 | 6:10 | 5:12 | 1:56 | 77.7% | 70.1% | 0.0031 |
| 2 | 9:14 | 8:54 | 0:32 | 7:37 | 7:19 | 0:48 | 82.5% | 82.2% | 0.0022 |
| 3 | 6:58 | 5:41 | 1:23 | 4:08 | 4:21 | 1:02 | 59.3% | 76.5% | 0.0045 |
| 4 | 9:31 | 8:22 | 1:07 | 6:14 | 7:00 | 1:12 | 65.5% | 83.7% | 0.0007 |
| 5 | 10:04 | 10:10 | 0:34 | 6:30 | 6:06 | 0:32 | 64.6% | 60.0% | 0.0027 |
| 6 | 9:41 | 8:21 | 1:14 | 6:15 | 5:06 | 0:33 | 64.5% | 61.1% | 0.0004 |
| 7 | 7:03 | 6:24 | 0:39 | 4:52 | 4:10 | 1:27 | 69.0% | 65.1% | 0.0014 |

The measured times reported in Table 2 show that for each of the considered tasks, the participants needed less time by using the SDArch system than the conventional one. The values of the relative user performance with respect to the average task completion time range from 59.3% to 77.7%. Moreover, calculated *p-values* for all tasks were statistically significant (i.e., < 0.05), which actually means that by using the SDArch system the participants needed significantly less time than by using the conventional system for all tasks.

Tables 3 and 4 show the same descriptive statistics for the number of mouse clicks and the number of window switches, respectively. For both of these metrics, regarding all the tasks, the relative performance of the participants when using the SDArch system with respect to the conventional system was less than 100%. In other words, the

participants performed less mouse clicks and window switches by using the SDArch system than by using the conventional one, regarding all considered tasks. In particular, a number of window switches for each task was significantly (more than two times) less with the SDArch system. Finally, *t-test* revealed that the difference between the results of the two control groups regarding the applied evaluation metrics is also statistically significant (*p-values* of all the tasks for both metrics are less than 0.05).

Table 3. Mouse clicks statistics

| Task | Conventional System | | | SDArch System | | | Relative Performance | | t-Test <i>p</i> (T<=t) |
|------|---------------------|--------|----------|---------------|--------|----------|----------------------|--------|---------------------------|
| | Avg. | Median | σ | Avg. | Median | σ | Avg. | Median | |
| 1 | 124.3 | 107 | 14.23 | 108.3 | 93 | 15.32 | 87.1% | 86.9% | 0.0003 |
| 2 | 137.2 | 118 | 15.31 | 109.2 | 102 | 9.30 | 79.6% | 86.4% | 0.0048 |
| 3 | 128.4 | 122 | 11.42 | 96.7 | 84 | 8.42 | 75.3% | 68.8% | 0.0001 |
| 4 | 141.9 | 124 | 19.21 | 112.0 | 104 | 11.52 | 78.9% | 83.8% | 0.0023 |
| 5 | 152.0 | 133 | 16.73 | 77.7 | 71 | 7.32 | 51.1% | 53.3% | 0.0022 |
| 6 | 144.6 | 136 | 10.82 | 82.6 | 77 | 6.80 | 57.1% | 56.6% | 0.0007 |
| 7 | 122.5 | 109 | 18.34 | 98.5 | 94 | 5.32 | 80.4% | 86.2% | 0.0026 |

The measured task completion times (Table 2), numbers of mouse clicks (Table 3) and numbers of window switches (Table 4) indicate that the user efficiency when using the SDArch system outperforms from the user efficiency when using the conventional system, with respect to these three applied evaluation metrics.

Table 4. Window switches statistics

| Task | Conventional System | | | SDArch System | | | Relative Performance | | t-test <i>p</i> (T<=t) |
|------|---------------------|--------|----------|---------------|--------|----------|----------------------|-------|---------------------------|
| | Avg. | Median | σ | Avg. | Median | σ | Median | Avg. | |
| 1 | 20.4 | 12 | 7.04 | 4.4 | 4 | 1.45 | 33.3% | 21.5% | 0.0041 |
| 2 | 17.3 | 15 | 3.22 | 6.2 | 4 | 1.82 | 26.6% | 35.8% | 0.0032 |
| 3 | 21.6 | 17 | 4.54 | 5.2 | 5 | 1.37 | 29.4% | 24.0% | 0.0009 |
| 4 | 23.2 | 18 | 5.21 | 6.6 | 5 | 2.32 | 27.7% | 28.4% | 0.0015 |
| 5 | 21.8 | 19 | 3.18 | 5.6 | 4 | 2.41 | 21.0% | 25.6% | 0.0023 |
| 6 | 22.4 | 19 | 3.71 | 6.3 | 5 | 1.62 | 26.3% | 28.1% | 0.0042 |
| 7 | 18.5 | 14 | 5.57 | 4.0 | 4 | 0.82 | 28.5% | 21.6% | 0.0028 |

The user satisfaction was the third evaluation criterion, besides the user effectiveness and efficiency, which we considered in the SDArch evaluation. The participants provided us their subjective feedback about the following user satisfaction dimensions: the system usefulness (S1-S5), ease-of-learning (S6) and ease-of-use (S7-S9). Table 4 shows average and median rating for all the nine statements. As all statements had been formulated as positive, the same rating value has the same interpretation for each of them. Statements S5 and S9, which express the overall satisfaction regarding the usefulness and ease-of-use respectively, were the two best-rated statements with an average rating of 4.8. The other statements were also rated as highly positive with average ratings ranging from 4.1 to 4.7.

Table 5. Subjective user feedback for the SDArch system

| Statement | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Average | 4.7 | 4.3 | 4.1 | 4.7 | 4.8 | 4.7 | 4.3 | 4.6 | 4.8 |
| Median | 5 | 4 | 4 | 5 | 5 | 4 | 4 | 4 | 5 |

Related Work

Extensive research has been carried out lately on the application of semantic web technologies to different aspects of the learning content authoring. There are numerous approaches such as (Duval et al., 2001) that attempt to semantically annotate learning content at different levels of granularity in order to make it easily discoverable and reusable. However, in most of these approaches, the semantic annotations of the leaning contents are stored in huge centralized repositories, which is contrary to the trends of the emerging Web of interlinked data (Berners-Lee et al., 2006). Two major problems with the centralized learning content repositories are the problem of restricted access and the weak integration with most frequently used authoring suites (e.g., MS Word and PowerPoint). There is still a

big lack of tools that would enable content authors to connect and search the existing learning content repositories directly from within the authoring suites. Besides semantic annotation and search, there are approaches (Dodero et al., 2005) that try to make use of ontologies and the semantic web protocols into collaborative authoring of learning content. Finally, there are approaches (Jovanović et al., 2006; Henze, 2005) that apply semantic web technologies to support dynamic assembly and personalization of learning content.

The use of social networking in e-learning has also started to attract a lot of attention recently. Terry Anderson has introduced the concept of 'educational social software', which he defines as networked tools that support and encourage individuals to learn together while retaining individual control over their time, space, presence, activity, identity and relationship (Anderson, 2005). The social networking tools have started to be applied to different aspects of the e-learning process. They are used to support a social constructivist approach to e-learning (Friensen et al., 2004), or more specifically, to support self-governed, problem-based and collaborative activities (Gillet et al., 2008). Moreover, the social networking has found its place in forming "optimal" student groups (Ounnas et al., 2008), thus helping teachers allocate students to groups based on a set of constraints. The use of social networking in e-learning has also initiated discussion about the extent to which social tools should be separated or integrated in learning management systems (LMS) (Gillet et al., 2005). Recent trends have showed moving e-learning beyond centralized and integrated LMS towards a variety of separate tools, which are used and managed by the students in relation to their self-governed work (Klamma et al., 2006).

In spite of the fact that semantic web technologies and social networking have already been applied to many aspects of e-learning, a full synergy of these two initiatives in the context of e-learning still needs to happen. To what extent e-learning and the learning content authoring can benefit from this synergy has to be answered through appropriate empirical studies. We believe that the empirical study that we conducted with SDArch represents a step in the right direction.

Conclusions

Although the use of semantic web technologies and social networking in the authoring of learning content has been studied extensively over recent years, the real benefit they brought to the authoring process is still unclear. Up to what extent these innovative technologies can improve the effectiveness and efficiency of authors in carrying out authoring tasks is a question that still seeks an answer. In our opinion, the right answer to this question can be obtained through an empirical study conducted with an authoring system that is featured by these technologies. In this paper, we presented the results of one such empirical study that we conducted to investigate the benefits of the novel, semantic document architecture (SDArch) with respect to the authoring of the course material. Since semantic web technologies and social networking are two pillars of SDArch, the benefits of SDArch naturally extend to them. Based on objective, quantitative measures of user effectiveness and efficiency, and the users' subjective feedback, we found that the use of semantic web technologies and social networking results in improvements of the authoring of the course material compared to the conventional authoring approach. In the future work, we plan to perform a new, long-term evaluation study with more participants and a larger document collection. We also plan to consider the application of some other evaluation metrics, in addition to those applied in this study, as well as to provide more comprehensive statistical analysis of collected data.

Acknowledgements

The research leading to the results presented in this paper has been partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement Number 247893 (TaToo – Tagging Tool based on a Semantic Discovery Framework project).

References

- Anderson, T. (2005). Distance learning – social software's killer app? *Paper presented at the Open Distance Learning Association of Australia*, Adelaide, Australia.

- Bastili, V.R., Gianluigi, C., & Rombach, D. (1994). The Goal Question Metrics Approach. *Encyclopedia of Software Engineering*, 1, 528-532.
- Berners-Lee, T., Hall, W., Hendler, J.A., O'Hara, K., Shadbolt, N., & Weitzner, D.J. (2006). A Framework for Web Science. *Foundation and Trends in Web Science*, 1(1), 1-130.
- Betty, C., & Allard, S. (2004). Technology and Human Issues in Reusing Learning Objects. *Journal of Interactive Media in Education*, 4(1), 1-25.
- Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340.
- Dodero, J.M., Diaz, P., Aedo, I., & Sarasa, A. (2005). Integrating Ontologies into the Collaborative Authoring of Learning Objects, *Int'l J. of Universal Computer Science*, 11(9), 1568-1576.
- Duval, E., Forte, E., Cardinaels, K., Verhoeven, B., Durm, R., Hendrikx, K., Forte, M., Ebel, N., Macowicz, M., Warkentyne, K., & Haenni, F. (2001). The ARIADNE Knowledge Pool System. In *Communications of the ACM*, 44(5), 73-78.
- Friesen, N. & Anderson, T. (2004). Interaction for lifelong learning. *British Journal of Educational Technology*, 35(6), 679-687.
- Gediga, G., & Hamborg, K.C. (1999). Isometric: A usability inventory supporting summative and formative evaluation of software systems. *Human-Computer Interaction*, 1, 1018-1022.
- Gillet, D., Sandy, E. H., Man, Y. C., & Christophe, S. (2008). Turning Web 2.0 Social Software into Versatile Collaborative Learning Solutions. In *ACHI '08 Proceedings of the First International Conference on Advances in Computer-Human Interaction*, 170-176.
- Gillet, D., Nguyen, A., & Rekik, Y. (2005). Collaborative Web-based Experimentation in Flexible Engineering Education," *IEEE Transactions on Education*, 48(4), 696-704.
- Henze, N. (2005). Personal Readers: Personalized Learning Object Readers for the Semantic Web. In *Proc. of the 12th Int'l Conf. on Artificial Intelligence in Education*, 274-281.
- Jovanović, J., Gašević, D., & Devedžić, V. (2006). Ontology-based Automatic Annotation of Learning Content. *Int'l J. on Semantic Web and Information Systems*, 2(2), 91-119.
- Klamma, R., Chatti, M. A., Duval, E., Fiedler, S., Hummel, H., Hvannberg, T., Kaibel, A., Kieslinger, B., Kravcik, M., Law, E., Scott, .P., Specht, M., Tattersall, C., & Vuorikari, R. (2006). Social Software for Professional Learning: Examples and Research Issues. In *Proceedings of ICALT*, 912-915.
- Nešić, S. (2009). Semantic Document Model to Enhance Data and Knowledge Interoperability. *Annals of Information Systems*, 6, 135-162.
- Nešić, S., Lelli, F., Jazyeri, M. & Gašević, D. (2009). Towards Efficient Document Content Sharing in Social Networks, *Proc. of the 2nd SoSEA workshop co-located with ESEC/FSE*, 1-8.
- Nešić, S., Jazyeri, M. & Gašević, D. (2010). Semantic Document Architecture for Desktop Data Integration and Management. In *Proceedings of the 22nd International Conference on Software Engineering and Knowledge Engineering*, SEKE, 73-78.
- Nešić, S., Crestani, F., Jazyeri, M., & Gašević, D. (2010). Search and Navigation in Semantically Integrated Document Collections. In *Proceedings of the 4th International Conference on Advances in Semantic Processing*, SEMAPRO. 55-66.
- Nielsen, J. (1993). Usability Engineering. Academic Press.
- Nielsen, J., Bevan, N., Barnum, C., Cockton, G., Spool, M., & Wixon, D. (2003). The "magic number 5": is it enough for web testing? In *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems*, 698-699.
- Ounnas, A., Davis, H., & Millard, D. (2008). A Framework for Semantic Group Formation. In *Proceedings of ICALT*, 34 – 38.
- Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargasvera, M., Motta, E., & Ciravegna, F. (2006). Semantic annotation for knowledge management: Requirements and survey of the state of the art. *Journal of Web Semantics*, 4(1), 14–28.
- Whittaker, S., Terveen, L., & Nardi, B.A. (2000). Let's stop pushing the envelope and start addressing it: A reference task agenda for HCI. *Human-Computer Interaction*, 15(2), 75-106.
- Zimmerman, D. (1997). Teachers Corner: A Note on Interpretation of the Paired Samples T-Test. *Journal of Educational and Behavioural Statistics*, 22(3), 349–360.

Designing Collaborative E-Learning Environments based upon Semantic Wiki: From Design Models to Application Scenarios

Yanyan Li¹, Mingkai Dong^{1,2} and Ronghuai Huang¹

¹Knowledge Science & Engineering Institute, School of Educational Technology, Beijing Normal University, China

//²Corporate Technology, Siemens Limited China, Beijing, China // liyy1114@gmail.com // dongmk@gmail.com // Huangrh@bnu.edu.cn

ABSTRACT

The knowledge society requires life-long learning and flexible learning environment that enables fast, just-in-time and relevant learning, aiding the development of communities of knowledge, linking learners and practitioners with experts. Based upon semantic wiki, a combination of wiki and Semantic Web technology, this paper designs and develops flexible e-learning environments for different application scenarios aiming to facilitate collaborative knowledge construction and maximize resource sharing and utilization. One application scenario is to support hybrid learning by deploying an online course platform and the first round of using has shown that the course platform can effectively facilitate and support students to fulfill task-driven learning in a more flexibly and friendly collaborative manner. The other application scenario is to build a teamwork platform for supporting collaborative e-research. After several months' trial, team members agree that the platform can well meet their collaborative research work demands with the advantage of quick, easy and convenient operating assistance. The kernel idea of the collaborative e-learning environments is to enable structural organization of resources with semantic association while providing diverse customized facilities.

Keywords

Collaborative knowledge construction, E-learning 2.0, Interactive query, Semantic wiki

Introduction

E-Learning has become one of the most popular teaching and learning methods by stretching the spatial and temporal barriers. Various e-learning systems have been developed in the past decade. Learning Management Systems are able to support online training with different levels of granularity and formalization, which focus on automation of some aspects of the design process, execution and assessment (Acqua, 2009). Learning Content Management System expects to provide standard-based content repositories that allow learners to capture, store, deliver, and manage learning resources (Goecks et al., 2002)(Jari et al., 2003). Furthermore, ubiquitous learning environment gains more and more attention in recent years, which aims to provide an interoperable, pervasive, and seamless learning architecture to connect, integrate, and share three major dimensions of learning resources: learning collaborators, learning contents, and learning services (Cheng et al., 2005)(Yang, 2006). However, almost the e-learning systems are self-independent and the learning resources are disordered, isolated, and heterogeneous, and there is no common overarching context for the available resources.

Virtual learning communities are cyberspaces in which individual and collaborative learning is implemented by groups of geographically dispersed learners and providers of knowledge to accomplish their goals of learning. Though there are no agreements on what constitutes a virtual learning community, it has gained widespread acceptance that virtual learning communities are knowledge based social entities where knowledge is the key to their success (Bhatt, 2001)(Malhotra, 2000). An important activity in a virtual learning community is the collaboration. Many virtual learning communities strive to attract new members or encourage members to learn and to contribute knowledge. Nevertheless, such collaboration environment is generally not supported by conventional learning environments (Hage et al., 2008).

E-learning 2.0 emerges inspired by the popularity of Web 2.0, which places increased emphasis on social learning and use of social software such as blogs, wikis, and etc. (Rosen, 2009). Conventional e-learning systems were based on instructional packets that were delivered to students using Internet technologies. The role of students consisted in learning from the reading and preparing assignments. By contrast, E-learning 2.0 is built around collaboration, which assumes that knowledge is socially constructed. Learning takes place through conversations about content and grounded interaction about problems and actions. Misanchuk et al. (2001) propose strategies focus on promoting communication, social interaction and participation to scaffold learning. Therefore, it is desirable to design and develop a sophisticated learning environment to achieve e-learning 2.0, encouraging learners' active involvement of

resource contribution, enabling convenient resources accessing and utilization, and facilitating better interaction and collaboration.

Semantic Web is the emerging landscape of new web technologies aiming at web-based heterogeneous resources that would be understandable and reusable by both humans and machines. Semantic knowledge is playing an increasing important role in order to have heterogeneous resources well organized and managed. Some researches have been done to utilize semantic web technologies to support e-learning (Kolovski et al., 2003)(Stojanovic et al., 2001) (Sampson et al., 2004). Wikis are well-known as online encyclopedias or websites that provide a vast source of information and allow each individual to contribute his own knowledge and experience on any topic (Clark, 2006). But a wiki is essentially a collection of Web sites connected via hyperlinks, which contain many hand-made, redundant, inconsistent lists and links and the meaning of its content is not machine-understood and machine-processable, so finding and comparing information from different pages is challenging and time-consuming.

By combining properties of wikis with Semantic Web technologies, semantic wikis emerged aiming to address problems due to the unstructured accumulated information of conventional wikis. The main idea is to make the inherent structure of a wiki – given by the strong linking between pages – accessible to machines beyond mere navigation. Semantic Wikis are structured, smart and accessible since they contain many lists that are computer-made, up-to-date and consistent. Although there are many researches on semantic wikis, much work focus on developing approaches to combining wikis with semantic technologies, which can be classified into two categories (Buffa, 2006). Most of the current projects on semantic wikis fall into the first category; i.e., they consider the use of wikis for ontologies, and the wikis become the front-end of the ontology maintenance systems, such as Platypus (Campanini, 2004), Rise (Decker, 2005), WikSAR (Aumueller et al., 2005), Semantic MediaWiki (Krötsch et al., 2007), etc. The second family of approaches focuses on the use of ontologies for wikis, and the typical semantic wikis include IkeWiki and SweetWiki (Buffa, 2006). So far, there have been many applications built upon different semantic wikis' engines. Most of the applications are on knowledge management and collaborative knowledge base construction (Oren, 2006) (Lange, 2008) (Happel, 2007), yet few applications on e-learning. With the conception of e-learning 2.0, semantic wiki is suitable to achieve e-learning 2.0 with the advantage of collaborative editing and structural resource organization. Nevertheless, search results on the semantic wiki platforms are displayed in a list of wiki pages, which is not intuitive for learners to understand easily. This appeal to reorganization of search results based on ontology category for quick browse and further study. Furthermore, the usual platforms only support users to view their own contributions and provide statistics on the wiki pages, but not consider the social relationship between users and can't recommend learning peers for collaborative learning and problem-solving. So learners usually feel isolated and are easily tended to become disengaged and inactive when they are outside of the social context.

Based on Semantic MediaWiki, this paper designs and develops Semantic Wiki-based Collaborative e-Learning Environments (SWiCLE) to better support learners collaborate to share, exchange and utilize learning resources in the sphere of e-learning 2.0. The kernel idea is to enable structural organization of learning resources with semantic association while providing diverse customized facilities, such as semantic search, multi-view filter, relevance-recommendation, etc.

Structuring Learning Resources

The kernel to structuring learning resources is to encapsulate learning materials with metadata descriptions and establish semantic linking between them. An ontology gives an explicit specification of a conceptualization with respect to the specific application domains by defining concepts and their properties, which lays the foundation to effectively organize and link resources with semantic relationship in an e-learning environment. Figure 1 illustrates an example of partial ontology within academic settings. As the figure shows, concepts include project, course, person, etc. These concepts' properties are omitted except for their relationship with other concepts. The relationship consists of two types. One type is the usual relationship, including *similar-to*, *prerequisite*, *is-a*, *part-of*, *related-to*, etc. The other type is the specific relationship between the concepts, such as *develop*, *participate*, and *study*.

In SWiCLE, any multimedia objects (e.g. a block of text, PDF documents, images) and real world objects (e.g. people, organizations and events) are encapsulated as learning objects (LO) with metadata description. The metadata for a generic learning object falls into two categories. One is content description indicating what the learning object

is about, which is the basic information used to select learning object. The other is context description that indicates when to present the learning object, which expresses the pedagogical information of a learning object as well as relationship with other learning objects. Currently, learning object metadata generation mainly depends on the manual annotation of resource contributors based on the pre-designed semantic templates. As a complement, the system can automatically extract some metadata by means of information extraction technology, which won't be discussed in this paper. Furthermore, a semantic link network (SLN) is a model to intuitively represent the semantic relationships between document fragments or documents (Li et al., 2009), and we adopt SLN to represent relationship between learning objects. Therefore, the resource organization structure of SWiCLE can be considered as two interconnected networks or spaces (see Figure 2). The knowledge space depicts the knowledge structure in SWiCLE, which is composed of Ontology and LO-SLNs. Ontology describes the generalized knowledge (concept and relationship between concepts) in an application domain, and a learning object in semantic link networks is an instance of a concept in the ontology. The learning objects belonging to the same concept constitute one LO-SLN, and the lines within one LO-SLN or across different LO-SLNs express the relationship between learning objects. The hyperspace is a network of hypertext pages with learning materials (traditional hyperspace). A learning object corresponds to a page that embodies the learning object's properties and related information (e.g. semantic relationship with other learning objects, annotations, comments).

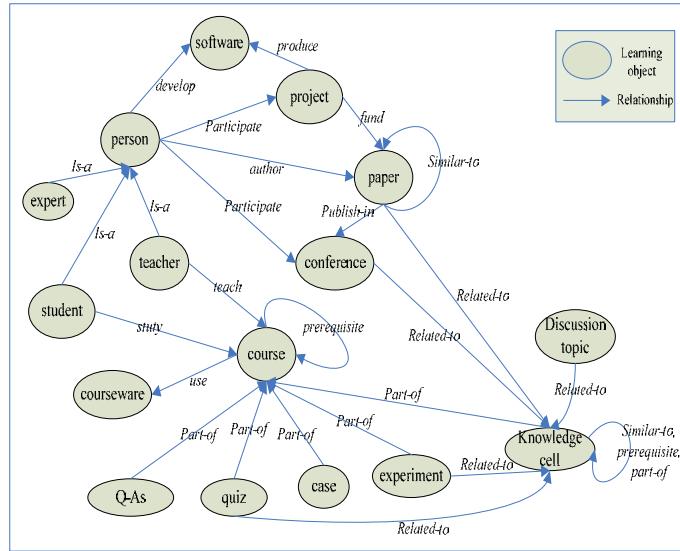


Figure 1. An example of partial ontology

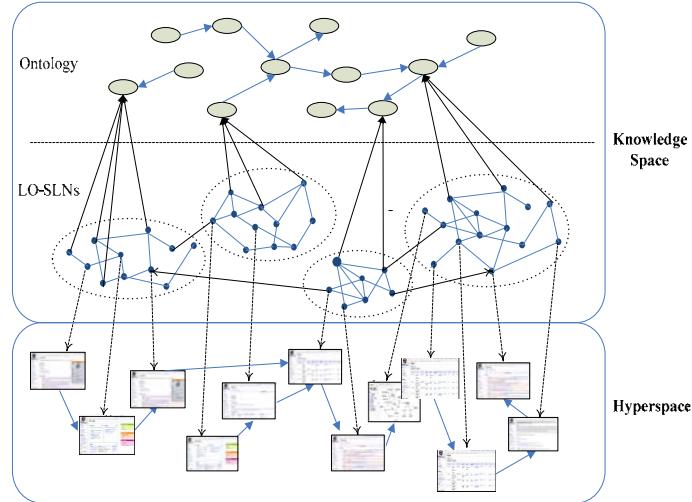


Figure 2. Conceptual model of resource organization

Semantic Search

In contrast to the traditional search, queries within semantic search lead to the focused search and quick location of the precise information and semantic related resources. SWiCLE provides the following three modes of semantic search.

Interactive Query

Generally, search results are ranked and sequenced in a descending relevance order, and learners have to browse the results one by one to determine which item is what they really needed. In this way, learners often feel tedious because of too many returned items. By contrast, we propose an interactive search that intends to help learners find what they need in a more easily and friendly way by displaying search results in an intuitive way.

When a learner inputs a query, he or she can search for the page of which title is exactly the same with the query keywords. If so the system will lead to the exact page, otherwise the system will return no matching result. The other situation is to search for related information for the given query. In such case, the system will firstly search for the matching instances in the repository and then return the corresponding categories (i.e. concepts in the ontology) that contain the instances. Each category is marked with a number to indicate the number of matching instances belonging to the category. Afterwards, learners can check which category is what they search for and then click the category. Accordingly, on the basis of matching instances belonged to the specific category, the system will search for more related instances and return to users as well. Taken the matching instances as the anchor ones, this step is to find semantic relevant instances. The simple approach for selecting the target instances for the one matching instance, purely based on the structure of the graph, is to collect the first N triples originated from the anchor instance, where N is the pre-defined traversal constraints. As for the case of two matching instances corresponding to the query, it is the key problem to find all the semantic association paths between the two instances so as to select relevant instances on the paths. The basic idea of the algorithm is to traverse the graph in a breadth-first order starting from the two instances. The following step is to rank the target instances in terms of their association weights. This can be accomplished by enabling a user to browse the ontology and mark a region (sub-graph) of nodes and properties of interest. If the discovery process finds some semantic association paths passing through these regions then they are considered relevant, while other associations are ranked lower or discarded. Furthermore, learners can click “all” to browse all the matching and related instances despite of the belonging category.

Script-based Advanced Query

This mode supports users to search for information with complex constraints, making dynamically generated query results available to users. By writing simple scripts in wiki pages according to syntax rules, users can obtain lists or tables composed of up-to-date information. For instance,

```
{ {#ask:  
[[Category: expert]]  
[[research interests::artificial intelligence]]  
|?name =  
|?affiliation address=  
|sort=name  
|order=desc  
}}
```

This query can be used to find out all experts whose research interest is “artificial intelligence”, and only two properties (i.e. name and affiliation address) of experts are displayed in the descending sequence of expert name. Alternatively, as shown in figure 3, users can use a dialogue box to define the query. That is, input the searching objects and properties in the left and right column respectively. This example is to find out all tasks that are assigned to one specific user and three properties are displayed in the searching results, including *status*, *start time*, and *end time*. Yet this search mode seems difficult for non-technical learners. We are trying to modify this based on visual template, and thus learners only need to set constraints by ticking checkbox.

Ask

| Query | | Additional printouts (optional) | |
|---|--|---|--|
| [[Category:task]] [[Performer::user:马韶茜]] | | <input type="checkbox"/> Status <input type="checkbox"/> Start time <input type="checkbox"/> End time | |
| Sort by column (optional) <input type="text" value="End time"/> Ascending <input type="button" value="▼"/> [Add sorting condition] | | <input type="button" value="Find results"/> <input type="button" value="Hide query"/> <input type="button" value="Querying help"/> | |
| Previous Results 1– 11 Next (20 50 100 250 500) | | <input type="checkbox"/> Status <input type="checkbox"/> Start time <input type="checkbox"/> End time | |
| JAVA学习-第三周任务 JAVA学习-第五周任务 JAVA学习-第六周任务 推荐英文论文阅读 下周论文阅读-20101014 第7周任务：GWT学习 第8周次任务：Select语句训练 第9周任务：JSP/Servlet学习 第10周任务：GATE的基本功能和JAPE规则学习 用户主题关系网络构建研究 | | Completed 18 September 2010 24 September 2010 Completed 30 September 2010 8 October 2010 Completed 9 October 2010 15 October 2010 Completed 9 October 2010 15 October 2010 Completed 14 October 2010 21 October 2010 Completed 16 October 2010 22 October 2010 Completed 24 October 2010 30 October 2010 Completed 1 November 2010 6 November 2010 To be processed 15 November 2010 21 November 2010 Processing 9 November 2010 30 December 2010 | |

Figure 3. The interface of script-based semantic search

Comparison Retrieval

Especially, a comparison retrieval mode is designed for learners to find relationships between two learning objects, e.g. the connections between different books or the commonalities of people. For the given keywords, firstly identify the matching learning objects in repository, then their corresponding properties description are analyzed to find the potential connecting terms, and finally the search results are displayed in two columns within one page where the top potential connecting terms are properly highlighted so that the relationships between the two learning objects can be easily identified. For example, input the query to search for two persons as “David W.Johnson” and “Edythe Holubec”. Their basic information, teaching courses, books and involved activities will be listed in two columns with the highlighted connecting terms that indicate their commonalities. More information refers to (Li et al., 2009).

Collaboration Structure Analysis for Relevance-Recomendation

In an open e-learning environment, learners usually feel isolated and are easily tended to become disengaged and inactive when they are outside of the social context of the classroom (Wu et al., 2002). Thus, it is important to providing learners with social support according to learners’ collaborative structure, which can be achieved by analyzing learners’ performances and interaction based on their activities on the platform. In this way, when a learner encounters a problem or asks for help, appropriate learning-companions can be recommended to enhance learners’ in-depth communication and learning. Also, the environment can recommend related resources that are contributed or browsed by other closely-related learning-peers.

In SWiCLE, learners can edit a specific page by means of four operations, including initiating a new page, adding text to an existing page, deleting text to an existing page, as well as restoring text to an existing page. As the editing history is all stored, we can compare different versions of a certain page based on text analysis technology to find out the operations executed by different learners.

Accordingly, several indicators can be computed for a learner, including participation, activity, and impact. Participation measures the involvement of a learner in the wiki pages. Activity indicates the activities a learner executed on wiki pages in a certain period of time. Impact is used to indicate a learner’s influence degree for his offered information. A learner’s impact is increased if other learners follow to edit the page initiated by the learner or view the page edited by the learner. On the contrary, his impact is decreased if his contribution text is deleted by other learners. These indicators can help to find out learners’ performances during collaborative learning process, and thus necessary assistance could be timely provided to enhance learning. Furthermore, we assume that it is more possible that two learners have similar interest if the number of pages they both edit is much bigger. So, mutuality is computed based on learners’ editing activities to measure the relationship between learners.

Participation:

$$P_i = \sqrt{\frac{N_i + V_{add}^i + V_{delete}^i + V_{restore}^i}{\sum_{f=1..u} (N_f + V_{add}^f + V_{delete}^f + V_{restore}^f)}}$$
 (1)

Where N_i denotes the number of pages initiated by the i th learner, V_{add}^i , V_{delete}^i , $V_{restore}^i$ denotes the number of pages whose text is added, deleted, and restored by the i th learner, respectively. u denotes the total number of learners.

Activity:

$$A_i = \frac{\alpha \bullet N_i + (1 - \alpha) \bullet P_i}{\Delta t_p} \times \frac{1}{(t - t_d) + \tau}$$
 (2)

Where N_i and P_i respectively represents the number of pages initiated and edited by the i th person during the period Δt_p , t is the current date and t_d is the date when the i th person edits the latest page. τ is the adjust parameter to avoid the denominator is zero, and it is initially assigned 1. α is parameter that adjusts the relative impact of user's operation.

Impact:

$$I_i = \alpha \bullet U_i + \beta \bullet V_i - \gamma \bullet D_i$$
 (3)

Where U_i denotes the number of learners who edit the page initiated by the i th learner, V_i denotes the number of editing times of a page initiated by the i th learner. D_i denotes the number of times the i th learner's contribution texts are deleted. α , β and γ are ajustable parameters and $\alpha + \beta + \gamma = 1$.

Mutuality:

$$M_{i,j} = \alpha \bullet |A_{i,j}| + (1 - \alpha) \bullet (\sum_{u \in A_{i,j}} \text{Min}(N_i^u, N_j^u))$$
 (4)

Where $A_{i,j}$ denotes the set of pages co-edited by both of the i th learner and j th learner, and $|A_{i,j}|$ is the number of pages co-edited by i th and j th learners. N_i^u (N_j^u) denotes the number of times the i th (j th) learner edits on the u th page that belongs to $A_{i,j}$. So the larger the value of mutuality is, the more close the relationship between the persons is, and vice versa.

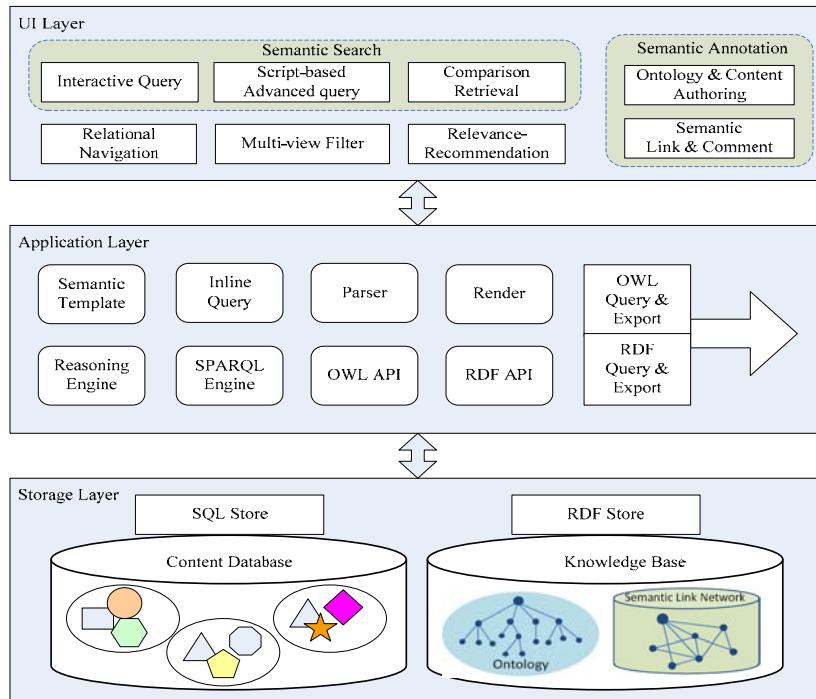


Figure 4. The architecture of SWiCLE

The Architecture of SWiCLE

Figure 4 illustrates the architecture of SWiCLE. As the figure shows, SWiCLE comprises three layers. The storage layer serves as the underlying learning repositories. The application layer consists of modules about processing and exchanging of structural information, which is accomplished by Semantic Web technologies. The UI layer provides various user interfaces for decentralized learners.

The storage layer is comprised of knowledge base and content database. Knowledge base consists of ontology and semantic link networks. The former is an explicit specification of a conceptualization with respect to the specific application domains, and the latter represents various semantic relationships between learning objects. Content database deposits learning materials authored or uploaded by learners and instructors.

In the application layer, semantic template comprises predefined parts of text into pages, and placeholders that are instantiated with user-supplied text when the template is included into a page. By simply adding typed links or attributes to the template text, the semantic template also allows the encapsulation of semantic annotation. Inline query enables editors to add dynamically created lists or tables to a page, thus making up-to-date query results available to readers who are not even aware of semantic queries. Parser is responsible for converting the text written by the user into learning objects. It parses the text for semantic annotations, layout directives, and links. Render takes charge of filling the page dynamically based on semantic templates, which determines the display layout delivered to learners. As extended from RAP (<http://sourceforge.net/projects/rdfapi-php>), the rest modules are used to create and store knowledge base. OWL Query & Export and RDF Query & Export give direct access to the formalized knowledge, which opens SWiCLE up to a wide range of external applications that will be able to use it as a background knowledge base.

Regarding the UI layer, several distinctive functionalities are provided. Semantic search comprises three search modes that are detailed in the former section. With the assistance of Ontology & Content authoring module, experts or administrators can construct and modify the domain ontology, and learners can create and edit learning objects along with properties via templates. Semantic linking allows learners to annotate links between learning objects via a special type of markup, while Comment empowers learners to evaluate learning objects with certain grading and remark. Relational navigation function offers additional information on the relation the semantic link describes, which changes the way content is presented based on semantic links and enables the content aggregation from different pages. This can include enriching pages by displaying of semantically related pages in a separate link box, displaying of information that can be derived from the underlying knowledge base, or even rendering its content of a page in a different manner that is more suitable for the context (e.g. multimedia content vs. text content). Multi-view filter offers faceted browsing to learners with advanced text search and filtering functionalities. With this function, property values of learning objects that occur more often can be grouped. Learners can select the values they would like to see by checking them, and then query results that do not hold the selected values will vanish. Relevance-recommendation function is to proactively recommend learning companions and related resources that are contributed or browsed by other closely-related learning-peers. This function is implemented based on analyzing learners' collaborative structure.

Applications

With an integrated, scalable and easy-to-use interface, SWiCLE serves as an entry point for learners to conveniently author, access, reuse and aggregate resources via diverse intelligent facilities, such as semantic search, relational navigation, relevance-recommendation, multi-view filter, etc.

Application Scenario 1: An online course for hybrid learning

As a preliminary application built on SWiCLE, we have developed an online course “Introduction to Artificial Intelligence”. By adopting task-driven collaborative learning strategy, the online course is flexibly designed to support hybrid learning. Figure 5 shows the semantic linked pages guided for relational navigation. During learning process, learners can browse learning content sequentially or click the marked blue text in the right flowing panel to study other content such as key concepts, previous task, etc. Meanwhile, learners can collaboratively edit each page

and comment on each page as well. To motivate students' collaboration, the final assignment of this course comprises three stages: individual design; mutual evaluation; and modification. Figure 6 shows one student's final assignment and mutual discussion with other students on the assignment. Additionally, as figure 7 illustrates, the course management module is especially designed for teachers allowing them to lookup students' activities, manage course resources and etc. Herein the interface shows students' contributions.



Figure 5. Semantic linked pages for relational navigation



Figure 6. Collaborative discussions on a student's final assignment

Twenty-two junior undergraduate students enrolled in the one-semester optional course. The course was carried out from March 2010 to June 2010 and took place 3 hour per week in a classroom. Besides listening in the classroom, students were required to finish one or two tasks for each learning unit through the online course platform. Each task is composed of several sequential learning steps.

An experimental study is conducted to understand students' perceived attitudes toward the use of semantic-wiki-based course platform. Furthermore, this study is to investigate the effects of course platform on quantity of collaborative editing and its influence on learning achievements. Thus, this study proposes two interesting research questions: 1) Students' attitudes toward the course platform; 2) Is the course platform able to improve students' learning results?



Figure 7. Course management for teachers

When the course ended up, a questionnaire to evaluate students' attitudes toward the course platform based on the semantic wiki was given to 22 students and 20 valid answer sheets were returned. The questionnaire included thirteen questions using a five-point Likert scale, which can be classified into four dimensions: (1) Perceived usefulness of the course platform; (2) Perceived ease of using the course platform; (3) Learning satisfaction of the course platform; (4) Willingness for future use. This questionnaire was originally developed by Davis (1989) and Azouaou et al. (2004), which is modified to evaluate how students come to accept and use PAMS 2.0 (Addison, 2010). Likewise, we modified the questionnaire based on the four dimensions. During questionnaire design process, we invite four experts whose major is psychology and computer to evaluate the validity of questions in terms of each dimension, and accordingly delete or modify some ambiguous or unsuitable questions. On the other hand, we asked students to complete the same questionnaire after three months, and then computed the test-retest reliability ($r=0.816$), which implies that the reliability of the questionnaire is sufficiently high.

Table 1 shows the result of this questionnaire. In terms of perceived usefulness, 100% students agree that course platform is useful to support their collaborative learning as indicated in question 1. In question 2, 85% of the students think that the course platform is useful to organize individual and group knowledge within a group. Likewise, in question 3, 95% of the students agree that the course platform is useful to share their thoughts with group members or all students. Nevertheless, regarding question 4, only 65% students think that the relational navigation is useful for learning. The reason for the unexpected results is that some students thought that the hierarchical task-step navigation mode is too deep to use, which lessen their enthusiasm to a certain degree. In terms of perceived ease of use, regarding question 5, 6, 7 and 8, most of students agree that it is easy to use the course platform. It's worthwhile to point that 25% students are not sure about the easiness of the course platform as indicated in question 7. The reason behind it is that these students seldom or never use web 2.0 systems before. In terms of learning satisfaction, 100% students are satisfied with the discussion function in the course platform. Likewise, regarding question 10 and 11, most of the students are satisfied with the course platform. In terms of willingness for future use, 80% students are willing to keep using the semantic-wiki-based platform in the future while the rest remain neutral. As for the last question, 75% students prefer using the course platform in other courses, 15% students do not decided, and 10% students are reluctant to use it in other courses. After interviewing with several students, we know that students are required to use different platform in different courses, which increases their working load and lessen their enthusiasm on using the platform for the formal learning. Yet they admit that they would like to use the platform for informal learning.

Table 1. Questionnaire results (N=20)

| | Questions | SA | A | U | D | SD | Mean |
|----------------------------|---|-----------|-----------|----------|----------|---------|------|
| Perceived usefulness | 1) I think the course platform is useful for supporting collaborative learning. | 5 25% | 15 75% | 0 0% | 0 0% | 0 % | 4.25 |
| | 2) I think the course platform is useful to organize individual or group knowledge in a group. | 10 50% | 7 35% | 3 15% | 0 0% | 0 0% | 4.35 |
| | 3) I think the course platform is useful to share individual thoughts with group members or all students. | 8 40% | 11 55% | 1 5% | 0 0% | 0 0% | 4.35 |
| | 4) I think the relational navigation function of the course platform is useful for learning. | 5 25% | 8 40% | 4 20% | 2 10% | 1 5% | 3.7 |
| Perceived ease of use | 5) I think the co-edit function of the course platform is very easy to use. | 6 30% | 10 50% | 3 15% | 1 5% | 0 0% | 4.05 |
| | 6) I think the discussion function of the course platform is very easy to use. | 7 35% | 10 50% | 2 10% | 1 5% | 0 0% | 4.15 |
| | 7) I think it is easy for me to learn how to use the course platform. | 5 25% | 10 50% | 4 20% | 1 5% | 0 0% | 3.95 |
| | 8) I think the course platform is easy to express individual thoughts. | 5 25% | 13 65% | 2 10% | 0 0% | 0 0% | 4.15 |
| Learning satisfaction | 9) I am satisfied with the discussion function in the course platform. | 5 25% | 15 75% | 0 0% | 0 0% | 0 0% | 4.25 |
| | 10) I am satisfied with peer interaction function in the course platform. | 5 25% | 13 65% | 2 10% | 0 0% | 0 0% | 4.15 |
| | 11) I think the course platform is suitable for collaborative learning. | 6 30% | 13 65% | 1 5% | 0 0% | 0 0% | 4.25 |
| Willingness for future use | 12) I would like to keep using the semantic-wiki-based platform in the future. | 6 30% | 10 50% | 4 20% | 0 0% | 0 0% | 4.1 |
| | 13) I would like to use course platform in other courses. | 5 25% | 10 50% | 3 15% | 2 10% | 0 0% | 3.9 |

Note. SA, strongly agree; A, agree; U, undecided; D, disagree; SD, strongly disagree.

Regarding the second research question, we perform significant correlation analysis between students' activities on the course platform and their learning achievement. Students' activities are mainly reflected by their editing behavior, and thus we adopt the formula $A_i = N_i + 2 * \sqrt{M_i - N_i}$ to compute the i th student's activities (A_i), where N_i denotes the number of pages initiated by the i th learner, M_i denotes the number of times of editing operation performed by the i th student. Students' learning achievement refers to their final grade that is obtained based on students' mutual marking and teacher's marking. The result of Spearman correlation analysis between students' behaviors and grades is shown in table 2. From the table, we can see that the two variables are moderately positive correlated ($R=0.548$, $p=0.012$), which denotes that students' activities on the course platform can affect their final grades to a certain degree.

Table 2. Correlation matrix for variables grade and behavior

Correlations

| | | grade | behavior |
|----------------|----------|-------------------------|----------|
| Spearman's rho | grade | Correlation Coefficient | 1.000 |
| | | Sig. (2-tailed) | .012 |
| | | N | 20 |
| | behavior | Correlation Coefficient | .548* |
| | | Sig. (2-tailed) | . |
| | | N | 20 |

*. Correlation is significant at the 0.05 level (2-tailed)

Application Scenario 2: A teamwork platform for collaborative e-Research

Scientific documents, research instruments, researchers' competence and cooperation among researchers are fundamental resources of scientific research. Research teams usually construct a platform or website to manage and

deliver their scientific resources, yet almost the platforms serve as the information propagation portals with simple documents management and communication functionalities. Based upon SWiCLE, We have constructed a sustainable teamwork platform to effectively support team members to capture, publish, share and manage explicit knowledge resources, to aid efficient scholarly communication and development of communities of knowledge, and to link novice with experienced researchers for better team building.



Figure 8. Ontology browser

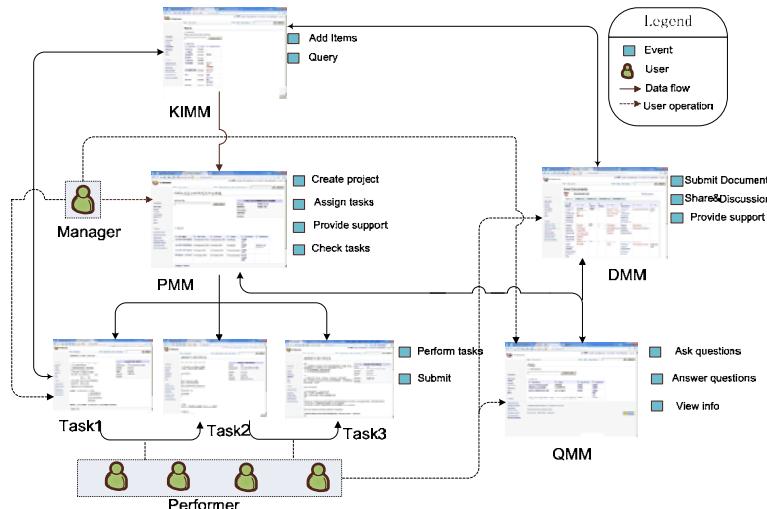


Figure 9. Collaborative working mode

As figure 8 shows, we have built an ontology for scientific research. The three colored parts from left to right displays the tree-structure categories (i.e. concepts), instances and corresponding properties, which allows users to easily browse ontology in an intuitive and friendly way. The teamwork ontology concepts comprise *project*, *task*, *people*, *document*, *question*, etc. Correspondingly, we design several modules to manage resources, such as project management module (PMM), document management module (DMM), knowledge item management module (KIMM), and question management module (QMM). Figure 9 illustrates the collaborative working mode. All the team members can share documents, propose questions or add knowledge items during working process. A manager can create a new project, assign relevant tasks to other team members, and submit some referential materials if needed. Then each team member can examine the projects or tasks allotted to him in his personal working space as illustrated in figure 10. In the personal working space, each user can also know about recent events and changes on the platform, manage his own documents, find all the questions he proposes for others or other team members

propose for him, as well as calendar setting and favorites management. Figure 11 is the interface to lookup timeline of the team. When the mouse moves upon any event, a small window will pop up automatically showing the event description. This function allows the users to have a quick view of the whole team's working planning and progress.

| Projects | | Members | Manager | Start time | End time |
|--------------------|------------------------|------------|---------|----------------|------------------|
| CSCL交互分析研究及平台搭建 | 马韶茜 王丽 | 李艳燕 张永和 | | 1 January 2011 | 31 December 2013 |
| 基于语义上下文的知识服务关键技术研究 | 冯凯 刘媛媛 王丽 马韶茜 | 李艳燕 | | | |

| Tasks | | Status | Priority | Start time | End time | Performer | Download |
|--------------|-----------------|--------|-------------------|-------------------|-----------|---------------------------|----------|
| JAVA学习-第三周任务 | Completed | High | 18 September 2010 | 24 September 2010 | 马韶茜 王丽 | | |
| JAVA学习-第五周任务 | Completed | High | 30 September 2010 | 8 October 2010 | 马韶茜 王丽 | Refderby.pdf MyDB.rar | |
| JAVA学习-第六周任务 | To be processed | High | 9 October 2010 | 15 October 2010 | 马韶茜 王丽 | | |
| 推荐英文论文阅读 | To be | Medium | 9 October 2010 | 15 October | 马韶茜 | +++using data mining as a | |

Figure 10. Personal working space

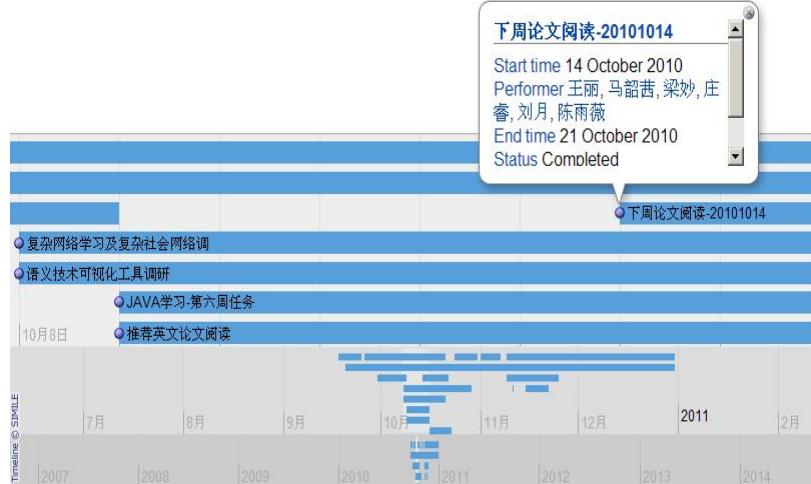


Figure 11. Team working timeline

To find out user's individual behaviors on the platform as well as mutual interaction among team members, we adopt the formulas mentioned before to conduct analysis in terms of users' participation, activity, impact and interaction for a period of 10 weeks. Figure 12 shows team members' participation statistics result. As the figure shows, there is a large gap between the users. The team member with the highest participation is the postgraduate who is responsible for the platform maintenance, and the team members with lower participation are almost undergraduates who less participated in the research. To further understand users' behaviors, figure 13 illustrates the changing trend of specific team members' activities over the period. There is one member who edits one page in the first week and no activities anymore because this student changed to other team. The user "yuanyuan liu" who is responsible for the platform maintenance are active during the period except one week, and the other two users' activities are relatively steady. Figure 14 shows team members' impact and mutual interaction, in which the node represents team member and the edge represents mutual interaction between team members. The bigger the node is, the higher impact of the member is. It is obvious that there are two members who have no interaction with others since they join the team very late; and besides, the team leader "yanyan li" has the highest impact, and then members "yuanyuan liu",

“yonghe zhang” as well as “shaoqian ma” have higher impact compared with others. Furthermore, the thickness of the edge denotes the interaction degree between team members. The thicker of the edge is, the closer interaction between members is and vice versa. Apparently, there is a clique within which the members have intensive interaction while the rest have fewer interactions with others. From figure 12 to 14, we can find that computation of indicators such as participation, activity, impact and mutuality can actually reflect users’ actions and interactions. This provides the basis to develop visualization modules that allow teachers or monitors to understand users’ performances on the platform in a more intuitive and convenient way.

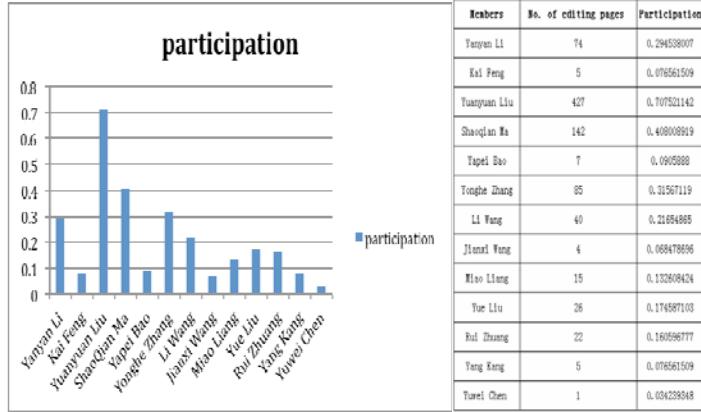


Figure 12. Participation statistics

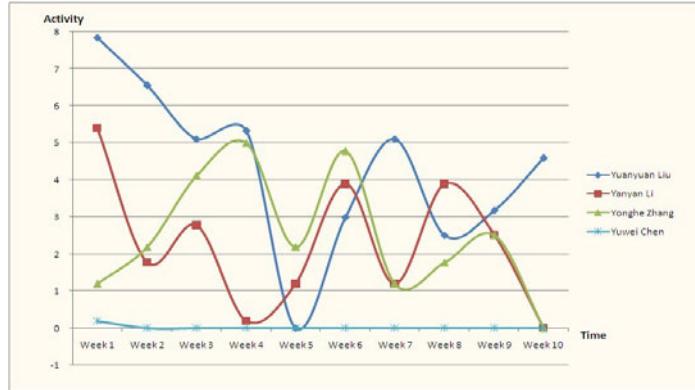


Figure 13. Changing trend of team members' activities

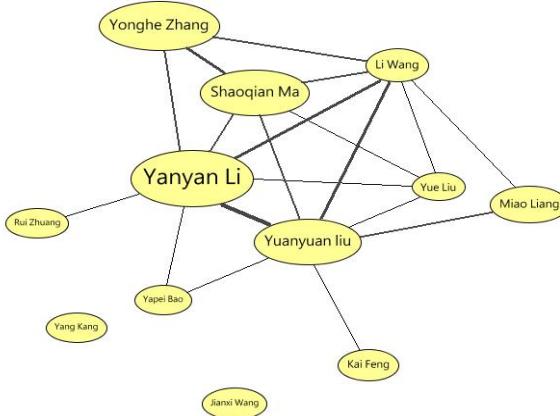


Figure 14. Team members' impact and mutual interaction

Discussion and Conclusion

Collaboration becomes an essential competency in the current knowledge society. In this study, a collaborative e-learning environment SWiCLE is designed based on semantic wiki to better support resource authoring, accessing, sharing and reusing via friendly facilities and to facilitate flexible collaborative learning as well. Compared with general e-learning platforms such as BlackBoard, Moodle, or other specific collaborative platforms (e.g. BBS, blog), Semantic wiki, a combination of wiki and Semantic Web technology, not only provides full functionalities of wikis for flexible interaction by allowing each individual to contribute his own knowledge and experience on any topic, but also enables structural organization as well as semantic association of resources.

As a preliminary exploratory application, we have developed an online course platform to support hybrid learning. Because it is an optional course, we don't know the students in advance. So we designed and developed the course based on task-driven mode without requirement analysis. The course comprises nine learning units and each unit comprises several tasks. To fulfill each task, students are required to follow the predetermined learning steps. Furthermore, several facilities are provided to support students' collaborative learning. Our initial thought is that the task-driven mode with collaborative knowledge building scaffold can lead students through learning process in an easier and effective way. When the course ended up, by means of questionnaire and interview, we learn that most of students are satisfied with the semantic-wiki-based course platform. Yet some students thought that this kind of hierarchical task-step navigation mode is too deep to use, which lessen their enthusiasm to a certain degree. On the other hand, the situation that some students are not familiar with wiki operation on the platform hinders them in learning process.

Regarding the collaborative teamwork platform, we conducted requirement analysis in advance and determined to design a lite platform to serve collaborative e-research within a team, such as task management, document management, discussion, etc. In order to organize various resources in a coherent and structured manner, we design an ontology and thus the increasingly added instances can be automatically semantic-linked according to the relationship between concepts defined in the ontology. Accordingly, each user own a personal working space in which he can conveniently view all the tasks assigned to him, all the documents shared with others, and all the questions proposed to him and he propose to others. As well, he can check his calendar and favorites easily. So far, almost each user agrees that the platform can satisfy his or her routine work in an easily and friendly manner, and especially, the platform can effectively support their collaborative research work.

Based upon the two applications, it can be seen that semantic-wiki-based platforms could provide diverse useful facilities for different application scenarios by enabling structural organization and semantic association of resources. Yet, it is worthwhile to note that ontology design plays an important role in resource structuring and semantic association. In addition, it is indispensable to conduct requirement analysis and then accordingly design different functionalities to cater for different applications.

Future work is to build a large-scale teacher training community based on SWiCLE to support life-long learning, whilst investigate the information flow and interaction content among users to analyze the community structure and user role, based on which to provide intelligent learning support for users in the community.

Acknowledgment

This research work is supported by National Natural Science Foundation of China (NSFC: 61075048), Fundamental Research Funds for the Central Universities (2009SD-9), and Arts Young Teacher Development Cultivation Project of Beijing Normal University.

References

- Acqua, L. (2009). A model for an adaptive e-learning environment. *Proceedings of the World Congress on Engineering and Computer Science (WCECS 2009)*, San Francisco, USA, 604-609.
- Aumueller, D., & Auer, S. (2005). *Towards a Semantic Wiki Experience – Desktop Integration and Interactivity in WikSAR*. Retrieved April 10, 2010, from http://dbs.uni-leipzig.de/person/david_aumueller.

- Azouaou, F., Chen, W., & Desmoulins, C. (2004). Semantic Annotation Tools for Learning Material. *Paper presented in International Workshop on Applications of Semantic Web Technologies for E-Learning*, August 23-26, Eindhoven.
- Bhatt, G. D. (2001). Knowledge Management in Organizations: Examining the Interaction between Technologies, Techniques, and People. *Journal of Knowledge Management*, 5 (1), 68-75.
- Buffa, M., Gandon, F., Ereto, G., Sander, P., & Faron, C. (2008). SweetWiki: A semantic wiki. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6, 84-97.
- Campanini, S.E., Castagna, P., & Tazzoli, R. (2004). Platypus Wiki: a Semantic Wiki Wiki Web. *Proceedings of the 1st Italian Semantic Web Workshop*, Ancona, Italy.
- Cheng, Z., Sun, S., Kansen, M., Huang, T., & He, A. (2005). A Personalized Ubiquitous Education Support Environment by Comparing Learning Instructional. *Proceedings of the 19th International Conference on Advanced Information Networking and Applications*, Tamkang University, Taiwan, 567-573.
- Clark, D. (2006). When informal learning works. *E.learning Age*, 8.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(1), 319-340.
- Decker, B., Ras, E., Rech, J., Klein, B., & Hoecht, C. (2005). Self-organized reuse of software engineering knowledge supported by semantic wikis. *Proceedings of the Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Galway, Ireland.
- Goecks, J., & Cosley, D. (2002). NuggetMine: Intelligent groupware for opportunistically sharing information nuggets. *Proceedings of the Intelligent User Interfaces Conference*, 87-94.
- Hage, H., & Aimur, E. (2008). Harnessing learner's collective intelligence: A Web2.0 approach to e-learning. *Lecture Notes in Computer Science*, 5091, 438-447.
- Happel, H., & Seedorf, S. (2007). Ontobrowse: A Semantic Wiki for Sharing Knowledge about Software Architectures. *Proceedings of Conference on Software Engineering and Knowledge Engineering*, retrieved May 1, 2011, from http://www.collabawue.de/dokumente/SK07-Seedorf_Happel_updated.pdf.
- Kolovski, V., & Galletly, J. (2003). Towards e-learning via the semantic web, *Proceedings of 2003 International Conference on Computer Systems and Technologies: e-Learning* (CompSysTech'2003), ACM, Rousse, 591-596.
- Lange, C. (2008). SWiM—a semantic wiki for mathematical knowledge management, The Semantic Web: Research And Applications. *Lecture Notes in Computer Science*, 5021, 832-837.
- Li, Y., & Dong, M. (2009). Building a Semantic Resource Space for Online Learning Community. *Proceedings of 2009 International Conference on Hybrid Learning* (ICHL 2009), 342-352.
- Malhotra, Y. (2000). Knowledge Management and Virtual Organizations. Hershey, PA: Idea Group.
- Misanchuk, M., & Anderson, T. (2001). *Building community in an online learning environment: communication, cooperation and collaboration*. Retrieved May 28, 2010, from <http://www.mtsu.edu/~itconf/proceed01/19.html>.
- Oren, E., Völkel, M., Breslin, J.G., & Decker, S. (2006). Semantic Wikis for Personal Knowledge Management. *Proceedings of the 17th International Conference on Database and Expert Systems Applications*, Krakow, Poland, 509-518.
- Rosen, A. (2009). *E-Learning 2.0: Proven practices and emerging technologies to achieve results*, New York: AMACOM.
- Sampson, D.G., Lytras, M.D., Wagner, G., & Diaz, P. (2004). Ontologies and the Semantic Web for E-learning. *Educational Technology & Society*, 7(4), 26-28.
- Stojanovic, L., Staab, S., & Studer, R. (2001). *E-learning based on the semantic web*. Retrieved April 10, 2010, from <http://www.uni-koblenz.de/~staab/Research/Publications/WebNet2001eLearningintheSemanticWeb.pdf>.
- Stroulia, E., & Jari, K. (2003). EduNuggets: an intelligent environment for managing and delivering multimedia education content. *Proceedings of the 8th International Conference on Intelligent User Interfaces*, Miami, Florida, USA, 303-306.
- Su, A.Y.S., Yang, S.J.H., Hwang, W.Y., & Zhang, J. (2010). A Web 2.0-based Collaborative Annotation System for Enhancing Knowledge Sharing in Collaborative Learning Environments. *Computer & Education*, 55, 752-766.
- Völkel, M., Krötzsch, M., Vrandecic, D., Haller, H., & Studer, R. (2006). Semantic Wikipedia. *Journal of Web Semantics*, 5, 251-261.
- Wu, A.S., Farrell, R., & Singley, M.K. (2002). Scaffolding Group learning in a Collaborative Networked Environment. *The International Conference on Computer-Supported Collaborative Learning*, Boulder, Colorado, USA.
- Yang, S. J. H. (2006). Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning. *Educational Technology & Society*, 9(1), 188-201.

A Workflow for Learning Objects Lifecycle and Reuse: Towards Evaluating Cost Effective Reuse

Demetrios G. Sampson and Panagiotis Zervas

Department of Digital Systems, University of Piraeus, Greece & Informatics and Telematics Institute, Centre for Research and Technology Hellas, Greece // sampson@unipi.gr // pzervas@iti.gr

ABSTRACT

Over the last decade Learning Objects (LOs) have gained a lot of attention as a common format for developing and sharing digital educational content in the field of technology-enhanced learning. The main advantage of LOs is considered to be their potential for component-based reuse in different learning settings supporting different learning activities. However, despite the importance of the concept of reuse and its potential benefits in digital educational content production and deployment, there are only sporadic efforts to study issues related to LOs reuse that would allow interested parties (such as people, organizations and initiatives) to assess the conditions for and eventually implement systematic LOs reuse within the context of learning activities design and development. This is a drawback in adopting the LOs paradigm towards reducing costs and effort. In this paper, we study existing efforts for the definition of the different steps involved during the LOs lifecycle, we identify the aspects of LOs reuse within the context of learning activities design and development, we propose a thorough workflow for LOs lifecycle that can support LOs reuse and enable us to define a set of metrics for cost effective LOs reuse, and discuss the cost effectiveness conditions in various use cases.

Keywords

Learning object, learning objects lifecycle, reuse workflow, cost metrics, cost effectiveness.

Introduction

Over the last decade Learning Objects (LOs) have gained a lot of attention as a common format for developing and sharing digital educational content in the field of technology-enhanced learning. Within the rich literature related to LOs, the issue of reuse appears to be an important one (Wiley, 2002; Polsani, 2003; McGreal, 2004; Caswell et al., 2008). This is mainly due to the fact that design and deployment process of high quality educational resources is very expensive, and therefore, any effort to reduce development costs is highly desirable (Zimmermann et al., 2006). However, despite the importance of the concept of reuse and its potential benefits in educational content production and deployment, there are only sporadic efforts to study issues related to LOs reuse that would allow interested parties (such as people, organizations and initiatives) to assess the conditions and eventually implement systematic LOs reuse within the context of learning activities design and development. This is a drawback towards the large scale adoption of the LOs paradigm aiming at reducing costs and effort.

In this paper, we study existing efforts for the definition of the different steps involved during the LOs lifecycle that can support LOs reuse and we identify their limitations. Based on the discussion of existing proposals, we propose a thorough workflow for LOs lifecycle that can support LOs reuse within the context of learning activities design and development. Finally, we use the proposed LOs lifecycle workflow to define a set of metrics so as to measure the cost effectiveness of LOs reuse and we extract recommendations that can facilitate interested parties to take more informed decisions about the potential benefits of LOs reuse.

Learning Objects Lifecycle and Reuse

What is Learning Objects Reuse?

The main arguments in favor of LOs reuse are twofold. On one hand, LO reuse is highlighted due to the anticipation of cost reductions in the design and development of educational resources while maintaining quality. This is based on the assumption that the more times a LO is reused in different learning settings the more cost effective that LO becomes. On the other hand, LO reuse can be an indicator for a high quality education resource. This is under the assumption that the more a LO is reused the more likely it is to be of high quality as more teachers and/or learners will have the opportunity to interact with it and provide feedback on its use and quality. However, despite the importance of the concept of LOs reuse, the technology-enhanced learning (TeL) community has not agreed to a

commonly accepted definition of the term ‘reuse’ resulting to multiple interpretations. The concept of LOs reuse, just as the concept of LOs, is presented in LOs literature in different ways as shown in Table 1.

Table 1. LOs Reuse Definitions

| Authors | Definitions |
|--|--|
| Wiley (2002, p. 12) | <i>“LOs can be used over and over again in similar contexts or in domains other than those for which they were designed”.</i> |
| Polsani (2003, p. 4) | <i>“A LO is predisposed to be reused in multiple instructional contexts”.</i> |
| Palmer & Richardson (2004, p. 5) | <i>“Reuse is the extent to which a LO can operate effectively for a variety of users in a variety of learning contexts over time in order to achieve the same or a different objective from that envisaged by its supplier”.</i> |
| Rensing et al. (2005, p. 4), Zimmermann et al. (2007, p. 49) | <i>“Reuse of LOs is any kind of use of existing LOs which are already used in a certain context for teaching or learning by trainers or learners in a new context to serve the same or a new purpose”.</i> |
| Colossus (2005, p. 1) | <i>“To reuse the LO with a different group of learners for which the LO was originally created”.</i> |
| Van Assche & Vuorikari (2006, p. 451) | <i>“Reuse is effective to the extent that a learning resource or any part of it can be fit into another learning resource or in another context for learning”.</i> |

Hence, based on the above definitions, we can conclude that the ability to reuse LOs includes the ability to reuse them in a different learning context and/or for a different targeted group and/or for the attainment of a different learning objective and/or for a different subject matter. Thus, one can note that the dimensions that affect the potential for LOs reuse are similar with the characteristics that define a learning activity (Beetham, 2007; Conole and Fill, 2005). According to Beetham (2007) a *learning activity* is a specific interaction of learner(s) with other(s) and with an environment (optionally involving resources, tools and services) that is carried out in response to a task orientated towards specific learning outcomes. Furthermore, according to Conole & Fill (2005) there are three (3) dimensions that constitute a learning activity:

- **The context** within which the activity occurs, this includes the subject matter (i.e., physics, geography, math, arts, etc.), the level of difficulty, the intended learning outcomes (i.e., recall, understand, etc.) and the environment within which the activity takes place (i.e., computer-based, lab-based, etc.).
- **The pedagogical approach** adopted (i.e., problem based learning, inquiry based learning, etc.)
- **The tasks** undertaken to achieve the intended learning outcomes. Tasks can be described by the type of task (i.e., reading, writing, viewing, etc.), the techniques used (i.e., presenting, discussing, arguing, etc.), associated tools and resources (i.e., computer, software, mobile devices, etc.), the interaction (i.e., class based, group based, etc.) and roles (teacher, learner, group leader, etc.) of those involved and the assessments (i.e., formative, summative) associated with the learning activity (Falconer et al, 2006).

Based on the above discussion and assuming that the pedagogical approach adopted can be considered as part of the context within which the activity occurs (Conole, 2007; Bailey et al., 2006; Weitl et al., 2004), in this paper we adopt the following definition for the concept of LO reuse: “Learning object reuse can be defined as the extent to which a Learning Object can be used in different digital or non digital learning activities, where a learning activity is defined as the interaction of learner(s) with other(s) and with a learning environment, which emerges as a result of performing a task within a particular learning context in order to achieve one or more learning objectives” (Sampson & Papanikou, 2009, p. 34)

Learning Objects Lifecycle

In order to study the process of LOs reuse, we need study the LOs lifecycle. In the literature there are some works that attempt to define the steps involved in the LOs lifecycle (Rensing et al., 2005; Collis & Strijker, 2004; Van Assche & Vuorikari, 2006). Most works study the LOs lifecycle in relation to the design and development of Learning Object Repositories (LOR). McGreal (2004) has defined LORs as systems that “enable users to locate, evaluate and manage learning objects through the use of “metadata”, namely descriptors or tags that systematically describe many aspects of a given learning object, from its technical to its pedagogical characteristics” (p. 3). First, Collis & Strijker (2004) argue that a LO can pass through six (6) different steps (following one another) during its lifecycle:

- **Obtaining:** the first step of the lifecycle is obtaining or creating a LO.
- **Labeling:** the LO created in the previous step is described with educational metadata.
- **Offering:** the LO is offered in a LOR so that other people can find it and retrieve it.
- **Selecting:** a user searches and selects from a LOR the LO that will suit the new needs.
- **Using:** after a LO is selected, it can be used either as it is in a new environment or modified in order to match the needs of the new environment within which the LO will be used.
- **Retaining:** after the use of the LO there are three possible choices, namely, the future use of the LO, its revision or its retraction from the LOR.

There are two main weaknesses in this proposal. First, in order for individual users to make use of existing LOs, they must be able to efficiently search for LOs and then evaluate the LOs returned as a result of that search, as to whether or not they are appropriate to be reused for meeting their specific expectations (Campbell, 2003). For that reason, in our work we propose that the step of “Selecting” should be explicitly identified as separate steps, namely, searching for appropriate LOs and selection of the most appropriate ones. Second, this proposal does not take into consideration the possibility of disaggregating a LO into its constituent parts and the selection of those suitable parts for the new learning activity (Colossus, 2005; Weitl et al., 2004). Therefore, if a LO is not reused in a learning activity as it is, then, two (2) more steps may be required, that is the modification and/or the aggregation with other LOs.

Another attempt to define the steps implemented in the LOs lifecycle was made by Rensing et al (2005) where the step of “Using” in (Collis & Strijker, 2004) is further analyzed. Considering both re-use (defined by Rensing et al (2005) as, any kind of use of existing LO, which are already used in a certain learning or teaching context) and repurposing (defined by Rensing et al., 2005) as, the modification of the LO in a way that suits a new learning or teaching context, which differs from the learning or teaching context that the LO was created for) this proposal identifies the extra steps of:

- **Modularization** of the LO, that is splitting the LO into several smaller LOs and selecting the appropriate ones.
- **Adaptation** of the LO, namely the modification of the LO with regard to at least one of its aspects (defined by Rensing et al (2005) as language, layout or terminology) to make it fit to a new learning or teaching context.
- **Aggregation** of the LO with other LOs to create a new one.

However, this proposal does not take into consideration issues that have important influence in time and cost of development, such as the selection of the appropriate LOs, the description of the LOs derived from the reuse process with metadata and the integration of the LO into the new learning or teaching context (Van Assche & Vuorikari, 2006).

The most complete effort for the explicit definition of the steps involved in the LOs lifecycle was made by Van Assche & Vuorikari (2006). The authors describe the LOs reuse in relation to a LO quality management policy and compared to the other two proposals, they add the following steps in the process of LOs lifecycle:

- **Approve**, where a LO before published in a LOR is reviewed (i.e., peer review) in order to ensure its high quality
- **Evaluate** that includes the criteria based on which the selection of suitable LOs for reuse is made
- **Integrate** that includes the technical (i.e., integration in a LMS) and/or pedagogical integration (expressed as the reshuffling the sequence of LOs in their proposal) of the LO into a new learning or teaching context.

Also, Van Assche & Vuorikari (2006) present the step of “Repurpose & Reuse” where the transformation of the LO takes place so that it can be reused in a new learning or teaching context. They argue that in this step the following actions may occur:

- **Disaggregation** of the LO into its constituent parts.
- **Aggregation** of the LO with other LOs.
- **Modification** of the LO content and/or of the sequence of the constituent parts of the LO.

Yet, in literature we can find more modification types that can be applied to a LO. These are divided into three (3) main dimensions (Zimmermann et al., 2006; Colossus, 2005; Duval & Hodgins, 2003):

- **Modifications to the LO layout/appearance**, when different LOs are combined to create a new LO, then modifications to LO appearance are needed or when different accessibility needs are addressed (i.e., people with

disabilities) then modifications to the display of the content is needed (i.e., white font and black background, so as to be accessible from low vision people)

- **Modifications to the LO content**, when different languages or terminology are addressed or when the sequence of the constituent parts of the LO is modified.
- **Modifications to the LO technical format** when different content delivery media and/or technology is addressed (i.e., mobile devices)

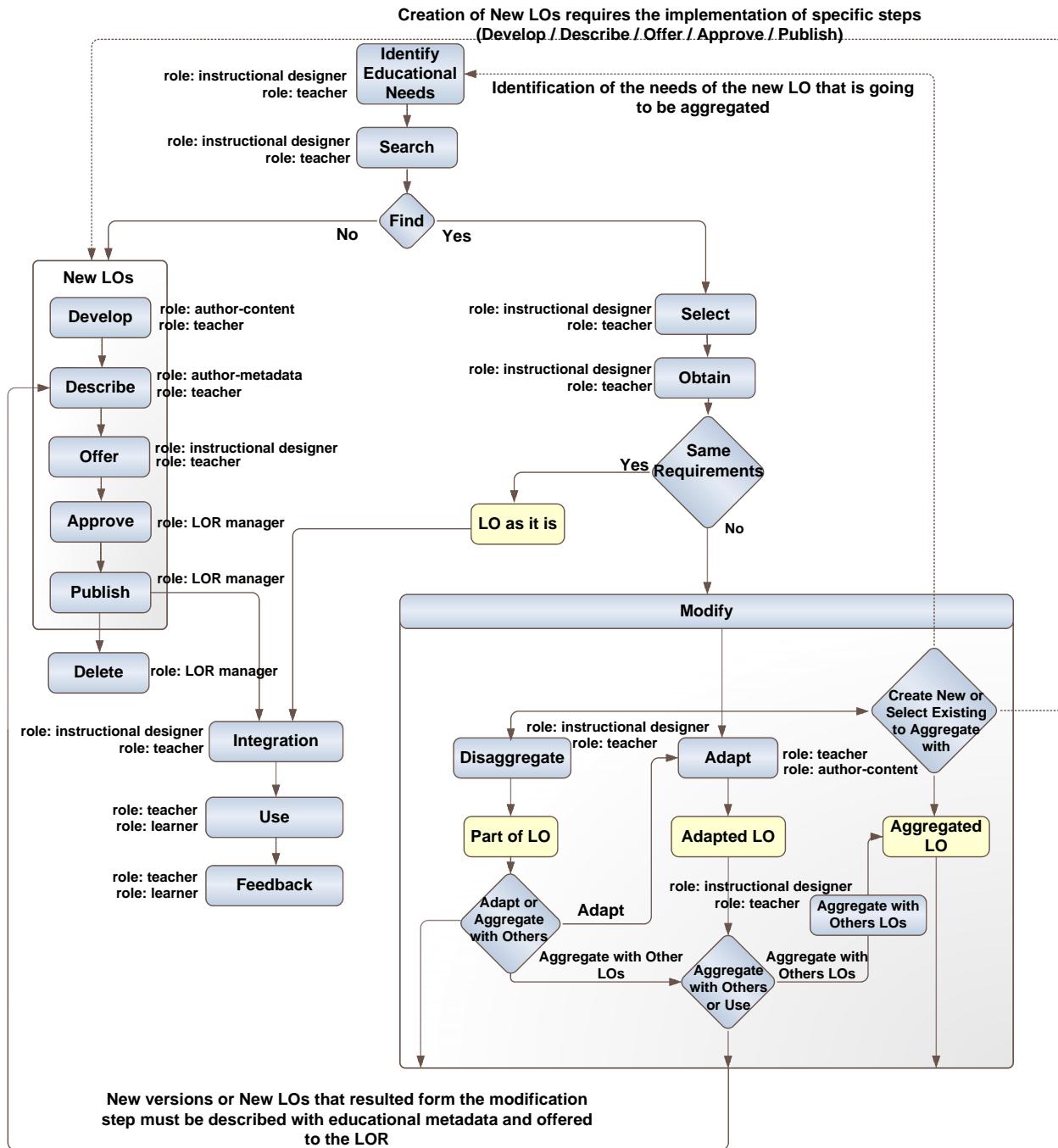


Figure 1. Learning Objects Lifecycle Workflow

Furthermore, none of the above approaches include in the LOs lifecycle the identification of needs that will lead to the selection of an appropriate LO or, if an appropriate LO does not exist, to the development of a new one. Identification of needs and intended learning outcomes are the first factors that influence the LO development

process (Palmer & Richardson, 2004). Finally, another important step not mentioned in the above approaches that encourages the LOs reuse in different learning or teaching contexts, is the step of LOs feedback. Feedback is defined as the process in which teachers and/or learners provide their advices/comments and/or ratings to a specific LO related to its use and quality (Weitl et al., 2004; Currier et al., 2004). Feedback is needed to support LO selection and maintain quality control. The feedback step could be integrated into the step of a LO's metadata characterization. However, it may include components such as rating that cannot be integrated in any of the IEEE LOM (IEEE LTSC, 2002) elements, so it is suggested to comprise an individual step.

Proposed Workflow for Learning Objects Lifecycle

Based on the discussion of existing proposals presented in the previous section, we propose a thorough workflow (including roles and their related functions) for LOs lifecycle in which LOs reuse is examined from the perspective of learning activities design and development. The proposed workflow assumes that the learning activities of an educational scenario have been already designed; we then begin with the phase of population learning activities with LOs. The participating roles in the workflow are the following:

- **Teachers:** Their role is twofold. On the one hand, they can develop new LOs to support their learning activities, possibly describe them with educational metadata and offer them to a LOR for future use by other users. On the other hand, they can re-use existing LOs (by applying modifications or not) to support learning activities towards the attainment of specific learning objectives.
- **Authors:** One can identify two (2) categories of authors, namely, the authors of educational content (role: author-content) and the authors of educational metadata (role: author-metadata). The authors of educational content (role: author-content) are responsible for the development of educational content in the form of LOs, ensuring that the produced LOs correspond to the learning objectives that were defined by the instructional designer. Authors may consist of sub-entities such as: (a) subject experts, who are responsible for developing learning content on a specific subject (i.e., astronomy, mathematics, biology), (b) graphic designers, who are responsible for developing the graphical elements of a LO, as well as, its look and feel and (c) technical developers, who make use of specific software tools in order to implement the desirable level of interaction among the LO and its user. The authors of educational metadata (role: author-metadata) are responsible for characterizing LOs with educational metadata.
- **Instructional Designers:** They define learning objectives and they design appropriate learning activities that will lead to the accomplishment of these objectives. They are responsible for designing and/or selecting appropriate LOs that will support the learning activities they wish to implement. Moreover, they facilitate authors to create and/or adapt LOs by providing advices regarding the instructional design of the LO, they can support the authors of metadata to describe LOs with educational metadata and they offer them to the LOR.
- **LOR Managers:** They are responsible for the LOR's policy, such as rights, terms of use and quality mechanisms. They approve and publish LOs to the LOR offered by the teachers and/or instructional designers. Finally, they are responsible for the possible retraction of LOs from the LOR.
- **Learners:** They are the final users of the LOs and the main participants in the learning activities. They also provide their feedback related to the use and quality of the LOs.

The functions that the proposed workflow includes are described below:

- **Identify Educational Needs:** The first function of the proposed workflow is the identification of educational needs. During this function the “role:instructional designers” or the “role:teachers” define the requirements that a LO must fulfill in order to be successfully used to support the learning activity they wish to implement. Therefore, the result of this function must be the explicit definition of the dimensions of the learning activity (as they have been defined in previous section) in which the LO will be used.
- **Search:** Before a LO is developed from scratch the “role:instructional designers” or the “role:teachers” searches the LOR (this can be one LOR or a federation of LORs) to examine if there is one or more existing LOs that fully or partly fulfill the requirements of the new learning activity (as defined in the previous function) and, therefore, they can be reused to a certain extent. Searching in a LOR includes searching based on criteria (fill in text fields or select a value from a vocabulary) that correspond to certain metadata elements and the return of one or more results which fulfill the search criteria. The result of this function is not a LO, but one or more metadata records that correspond to the search criteria. If the search results do not return a LO that fulfill these requirements, then the “role:instructional designers” can inform the “role:authors-content” to proceed to the

function of “Develop” a new LO. Alternatively, the “role:teachers” can proceed to the function of “Develop” a new LO. Otherwise, the “role:instructional designers” or the “role:teachers” proceed to the function of “Select”.

- **Develop:** At this function the “role:authors-content” or the “role:teachers” develop a new LO to support the learning activity with the requirements defined by the roles that participate to the function “Identify Educational Needs”.
- **Describe:** At this function the LO developed in the previous function is described with educational metadata following either IEEE LOM or an application profile created to serve specific needs. The “role:authors-metadata” characterize with educational metadata the LO developed by the “role:authors-content” or the “role:teachers” characterize with educational metadata the LO that they have developed during the previous function.
- **Offer:** The LO that has been already described with metadata in the previous function is offered to the LOR by the “role:instructional designers” or the “role:teachers”, so that other users can use it.
- **Approve:** Before a LO is published to the LOR and made available to its users, it may be reviewed (according to the LOR policy) by the “role:LOR managers”(i.e., peer review) in order to ensure its quality.
- **Publish:** Since a LO has been described with educational metadata and considered to be suitable for use, it can be made available (with or without usage restrictions or cost) by the “role:LOR managers” to other users of the LOR.
- **Select:** The “role:instructional designers” or the “role:teachers” in this function should evaluate the LOs returned as a result of the function “Search” in order to select the one that satisfies to a certain extent the requirements of their learning activity. The fundamental criterion that should affect the decision of LO selection must be the requirements defined in the function “Identify Educational Needs”. If a LO fulfills those requirements, then it can be reused as it is. Otherwise the LO must be modified in order to meet the specific requirements of the learning activity in hand. Other criteria that influence the decision of LO selection are comments made by other roles (role:learners and/or role:teachers), evaluations (i.e., peer review) of the LO or number of users downloaded the LO. A LO selection may be also based on copyright restrictions or cost.
- **Obtain:** Since the appropriate LO has been selected, the “role:instructional designers” or the “role:teachers” can obtain it. This sometimes requires usage permission by the owner of the LO or payment. Provided that LO fulfills the requirements of the new learning activity at the function of “Select”, then the “role:instructional designers” or the “role:teachers” can reuse the LO directly after integrating it into their learning activity. Otherwise they must go the function of “Modify”.
- **Modify:** Often, direct reuse of a LO is not feasible because it does not match the requirements of the learning activity that it will be used, as a result the following sub-functions may occur:
 - *Disaggregate*: In this sub-function, the “role:instructional designers” or the “role:teachers” decompose the LO into its constituent parts and those parts that match the requirements of the new learning activity are identified. The disaggregated LO constitutes a new LO. However, this LO may not be suitable as it is to cover completely the requirements of the new learning activity. Therefore the “role:instructional designers” should inform the “role:authors-content” to proceed to the function of “Adapt” an existing LO. Alternatively, the “role:teachers” can proceed to the function of “Adapt” an existing LO.
 - *Adapt*: In this sub- function, the “role:authors-content” or the “role:teachers” modify the LO, so as to fit to the requirements of the new learning activity. Adaptations may occur in the three (3) different dimensions that were defined in the previous section, namely, adaptations to the LO layout/appearance, adaptations to the LO content and adaptations to the LO technical format.
 - *Aggregate with other LOs*: In this sub-function, the “role:instructional designers” or the “role:teachers” aggregate the LO with other(s) LO(s) and thus a new LO is created. The LOs used for aggregation may result from the selection through the LOR or may be new LOs developed from scratch. When existing LOs are used, then their disaggregation or adaptation may be required.
- **Integrate:** At this function the “role:instructional designers” or the “role:teachers” integrate the LO into the environment that supports the learning activity in hand.
- **Use:** At this function the LO is used in a specific learning activity by the “role:learners” and/or the “role:teachers” towards the attainment of specific learning objectives.
- **Feedback:** In order for the LOs to be retrieved and used effectively in different learning activities more information are required about how they were used in practice, beyond the information derived by their educational metadata records provided by the “role:authors-metadata” or the “role:teacher”. A number of techniques are used in order for “role:learners” and/or “role:teachers” to provide feedback in the LOs of a LOR. The most commonly used techniques are comments (referring to the context of use of the LO and its usefulness)

- and ratings (the use of star ratings and/or hit counters that illustrate the number of downloads of a certain LO give a good indication of users' impression about the LO) (Gehringer et al., 2007; Kay & Knaack, 2007)
- **Delete:** The “role:LOR managers”, who are responsible for publishing a LO may decide that the LO must be retracted and, therefore, removed from the LOR, under certain circumstances.

Metrics for Cost Effectiveness of LOs Reuse

In this section, we use the proposed workflow of LOs lifecycle, so as to define metrics for cost effective LOs reuse, which can facilitate interested parties (people, organizations and initiatives) to assess the cost for systematic LOs reuse. Despite the importance of the concept of LOs reuse and its potential benefits, it seems that there are not proposed metrics for measuring the cost of LOs reuse, so as to enable us to perform a cost-benefit analysis. For this purpose, we propose to identify and adapt relevant cost metrics as in the field of software engineering.

Related Work

In the field of software engineering, reuse is considered as a very important factor for productivity and quality of software systems. As a result, a number of methods have been developed to measure the cost effectiveness of software code reuse (Frakes & Terry, 1996). Component-based Software Development (CBSD) is commonly accepted as a cost effective approach, as it emphasizes on the creation of software systems using reusable components (Washizaki et al., 2003). However, although software components reuse promises reduction in the development cost and time, as well as benefits in productivity and quality, its application in practice does not necessarily ensure that these benefits can be achieved. Therefore, appropriate metrics and models have been proposed as tools to measure and assess the impact of reuse (Hafeh et al., 2002).

Within this context, Poulin et al. (1993) described a set of cost metrics for software components re-use used by the IBM company (<http://www.ibm.com>) that are the most commonly used mainly because they are simple to understand and easy to calculate during the software development process (Mascena et al., 2005). Theis main cost metrics are:

- **Relative Cost of Reuse (RCR),** which is defined as the cost for reusing a software component divided by the cost normally incurred to develop it for one-time use,
- **Relative Cost of Writing Reusable Software (RCWR),** which is defined as the cost for developing a reusable software component divided by the cost of developing it for one-time use.

These metrics can be used as input in a return on investment model (ROI), upon which managers may rely their business decisions.

In the TeL literature, there are some works that have applied metrics from the software engineering field for the purpose of measuring the potential reusability of learning objects. Cuadrado & Sicilia (2005) explores the possibility of using existing object oriented design metrics proposed by Chidamber & Kemerer (1994) and adapting them to the LO domain, so as to measure the complexity of individual LOs internal structure and consequently assess their potential reusability. Cervera et al (2009) have also adapted these metrics in their study to measure potential reusability and quality of individual LOs by means of correlation between these metrics and their metadata. Finally, Mat Noor et al (2009) applied the metrics proposed by Cuadrado & Sicilia (2005), so as to measure the potential reusability of individual LOs selected from existing LORs (such as MERLOT and SMETE). However, these works assess only the potential reusability of individual LOs and they do not propose metrics for measuring whether the process of LOs reuse is cost effective in practice. In order to achieve that, we should be able to perform a cost-benefit analysis within a well-defined workflow of the LOs lifecycle (this has been defined in previous section), where cost variables can be assigned for each function of the workflow and metrics for cost effective LOs reuse can be defined.

Proposed Metrics

In this section, we assign cost variables that correspond to all different functions of the proposed workflow of LOs lifecycle, so as a cost-benefit analysis to be feasible. Table 2 presents these variables.

Table 2. Identified Costs of the Proposed LOs lifecycle Workflow

| LOs Lifecycle Workflow Functions | Cost Variable |
|-------------------------------------|--------------------|
| Identification of Educational Needs | C_{needs} |
| Search | C_{search} |
| Selection | C_{select} |
| Obtain | C_{obtain} |
| Disaggregation | $C_{disaggregate}$ |
| Adaptation | C_{adapt} |
| Aggregation with Other | $C_{aggregate}$ |
| Integration | $C_{integrate}$ |
| Feedback | $C_{feedback}$ |
| Description | $C_{metadata}$ |
| Offer | C_{offer} |
| Approval | $C_{approve}$ |
| Publish | $C_{publish}$ |
| Development | $C_{develop}$ |

Next, we present a set of metrics that can facilitate measuring the cost effectiveness of LOs reuse.

Cost to Create a Single Non-Reusable LO (C2CNRLO)

This metric is defined as the cost needed to develop a non-reusable LO from scratch. According to the proposed LOs lifecycle workflow, the functions that are needed to develop a single non-reusable LO are: a) identify educational needs, b) develop, c) integrate and d) feedback. As a result, C2CNRLO metric can be calculated using the following formula:

$$C2CNRLO = C_{needs} + C_{develop} + C_{integrate} + C_{feedback}$$

Additional Cost for Reusable LO (ADC4RLO)

This metric is defined as the additional cost needed to create a reusable LO. According to the proposed LOs lifecycle workflow, the additional functions that are needed to develop a reusable LO are: a) describe, b) offer, c) approve and d) publish. As a result, ADC4RLO metric can be calculated using the following formula:

$$ADC4RLO = C_{metadata} + C_{offer} + C_{approve} + C_{publish}$$

We should mention here that ADC4RLO takes its maximum value if the particular LO is reused only once. Provided that the particular LO is frequently re-used, then ADC4RLO could be reduced to practically zero

Cost to Create a Single Reusable LO (C2CRLO)

This metric is defined as the cost needed to create a reusable LO from scratch. According to the proposed LOs lifecycle workflow, the cost needed to create a reusable LO (C2CRLO), includes the cost needed to create a non-reusable LO (C2CNRLO), as well as the additional cost needed to create a reusable LO (ADC4RLO). As a result, C2CRLO metric can be calculated using the following formula:

$$C2CRLO = C2CNRLO + ADC4RLO$$

Cost to Create a Sequence of LOs within a new Learning Activity (C2CLO)

This metric is defined as the cost needed to create from scratch non-reusable LOs (C2CNRLO) and/or reusable LOs (C2CRLO) for the needs of a new learning activity. This metric can be calculated as follows:

$$C2CLO = \sum_{i=1}^{K1} C2CNRLO_i + \sum_{i=1}^{K2} C2CRLO_i$$

Where:

- (K_1) is the number of non-reusable LOs developed for the purpose of the new learning activity
- (K_2) is the number of reusable LOs developed for the purpose of the new learning activity

Cost to Reuse a Single LO within a new Learning Activity (C2RLO)

This metric is defined as the cost needed to reuse a LO (with or without modifications). According to the proposed LOs lifecycle workflow, we should examine two (2) cases:

- *Cost to reuse a LO without modifications in the new learning activity (C2RLO_{AsIs})*: when a LO is reused without modifications in a new learning activity, then the functions that are implemented based on the proposed LOs lifecycle workflow are the following: a) identify educational needs, b) search, c) select, d) obtain, e) integrate and f) feedback. As a result, C2RLO_{AsIs} metric can be calculated using the following formula:

$$C2RLO_{AsIs} = C_{needs} + C_{search} + C_{select} + C_{obtain} + C_{integrate} + C_{feedback} = (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain}$$

- *Cost to reuse a LO after modifications in the new learning activity (C2RLO_{modify})*: when a LO is reused after modifications in a new learning activity, then the functions that are implemented based on the proposed LOs lifecycle workflow are the following: a) identify educational needs, b) search, c) select, d) obtain e) disaggregate f) adapt, g) aggregate with others, h) describe, i) offer, j) approve, k) publish, l) integrate and m) feedback,. In this case except the additional functions (in relation to the case of reusing a LO without modifications) that may emerge due to LO modification (namely, disaggregate, adapt, aggregate with other LOs), the functions of description, offer, approval and publish to the LOR have been added, since it is most likely that a modified LO needs to have its educational metadata updated and it must be offered to the LOR as a new LO in order to be available to other users. Consequently, within the context of calculating the Cost to Reuse a LO after modifications, we should examine three complementary (3) cases:

- Cost to reuse after disaggregation (C2RLO_{disaggregate}): this metric can be calculated using the following formula:

$$C2RLO_{disaggregate} = C_{needs} + C_{search} + C_{select} + C_{obtain} + C_{disaggregate} + C_{metadata} + C_{offer} + C_{apprrove} + C_{publish} + C_{integrate} + C_{feedback} = (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain} + C_{disaggregate}$$

- Cost to reuse after adaptation (C2RLO_{adapt}): this metric can be calculated using the following formula:

$$C2RLO_{adapt} = C_{needs} + C_{search} + C_{select} + C_{obtain} + C_{adapt} + C_{metadata} + C_{offer} + C_{apprrove} + C_{publish} + C_{integrate} + C_{feedback} = (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain} + C_{adapt}$$

- Cost to reuse after aggregation with other LOs (C2RLO_{aggregate}): this metric can be calculated using the following formula:

$$C2RLO_{aggregate} = C_{needs} + C_{search} + C_{select} + C_{obtain} + C_{aggregate} + C_{metadata} + C_{offer} + C_{apprrove} + C_{publish} + C_{integrate} + C_{feedback} = (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain} + C_{aggregate}$$

As a result, the cost to reuse a single LO within a new learning activity could be equal to the following minimum and maximum values:

$$C2RLO = \begin{cases} (\min) (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain} = C2RLO_{AsIs} \\ (\max) (C2CRLO - C_{develop}) + C_{search} + C_{select} + C_{obtain} + (C_{aggregate} + C_{disaggregate} + C_{adapt}) = C2RLO_{modify} \end{cases}$$

The total cost of reusing LOs in a new learning activity can be calculated as follows:

$$C2RLO = \sum_{i=1}^{M1} C2RLO_{AsIs\ LOi} + \sum_{i=1}^{M2} C2RLO_{modify\ LOi}$$

Where:

- (M_1) is the number of LOs reused without modifications for the purpose of the new learning activity
- (M_2) is the number of LOs reused after modifications for the purpose of the new learning activity

Cost Benefit due to Reuse LO (CB2RLO)

This metric is defined as the total cost benefit that derives from the total cost of creating a sequence of non-reusable LOs and/or reusable LOs minus the cost of reusing LOs (with or without modifications) for the same learning activity. As a result, CB2RLO metric can be calculated using the following formula:

$$CB2RLO = C2CLO - C2RLO$$

Discussion

Based on the proposed metrics for measuring the cost effectiveness of LOs reuse, we can discuss the conditions of different cases, in which LOs reuse can be considered as cost effective. For this purpose, we examine the Cost Benefit due to Reuse (CB2RLO) metric, which should have a positive value, so as to consider that the LOs reuse is cost effective. This means that the following formula should be valid:

$$CB2RLO = C2CLO - C2RLO > 0 \Rightarrow C2CLO > C2RLO \Rightarrow$$

$$\sum_{i=1}^{K_1} C2CNRLO_i + \sum_{i=1}^{K_2} C2CRLO_i > \sum_{i=1}^{M_1} C2RLO_{AsIsLOi} + \sum_{i=1}^{M_2} C2RLO_{modifyLOi} \quad (1)$$

Assuming that: $K_1+K_2=M_1+M_2$

From the above formula, we can consider the following four (4) cases:

1. *The learning activity can be designed with non-reusable LOs that are developed from scratch or by reusing LOs without any modification:* for this case, formula (1) is transformed as follows:

$$\sum_{i=1}^K C2CNRLO_i > \sum_{i=1}^K C2RLO_{AsIsLOi}. \text{ By analyzing this formula, we get the following result:}$$

$$\sum_{i=1}^K C_{develop} > \sum_{i=1}^K (C_{search} + C_{select} + C_{obtain}).$$

2. *The learning activity can be designed with non-reusable LOs that are developed from scratch or by reusing LOs which have been all modified:* for this case, formula (1) is transformed as follows:

$$\sum_{i=1}^K C2CNRLO_i > \sum_{i=1}^K C2RLO_{modifyLOi}. \text{ By analyzing this formula, we get the following result:}$$

$$\sum_{i=1}^K C_{develop} > \sum_{i=1}^K (C_{search} + C_{select} + C_{obtain} + C_{disaggregate} + C_{aggregate} + C_{adapt} + ADC4RLO).$$

3. *The learning activity can be designed with reusable LOs that are developed from scratch or by reusing LOs without any modification:* for this case, formula (1) is transformed as follows:

$$\sum_{i=1}^K C2CRLO_i > \sum_{i=1}^K C2RLO_{AsIsLOi}. \text{ By analyzing this formula, we get the following result:}$$

$$\sum_{i=1}^K (C_{develop} + ADC4RLO) > \sum_{i=1}^K (C_{search} + C_{select} + C_{obtain}).$$

4. *The learning activity can be designed with reusable LOs that are developed from scratch or by reusing LOs which have been all modified:* for this case, formula (1) is transformed as follows:

$$\sum_{i=1}^K C2CRLO_i > \sum_{i=1}^K C2RLO_{modifyLOi}. \text{ By analyzing this formula, we get the following result:}$$

$$\sum_{i=1}^K C_{develop} > \sum_{i=1}^K (C_{search} + C_{select} + C_{obtain} + C_{disaggregate} + C_{aggregate} + C_{adapt}).$$

If we group the costs, $C_{search} + C_{select} + C_{obtain}$ and consider them as a total cost for searching and obtaining LOs from a typical LOR and if we also group the costs $C_{disaggregate} + C_{aggregate} + C_{adapt}$ and consider them as a total cost for modifying an existing LO then from the formulas described above, we can conclude the following:

- **Case 1:** The process of reusing a sequence of LOs (without any modifications) for a new learning activity is cost effective only if the sum of the costs to search and obtain them from a LOR is lower than the sum of the costs to develop them (as non-reusable LOs) from the scratch.
- **Case 2:** The process of reusing a sequence of LOs (with modifications) for a new learning activity is cost effective only if the sum of the costs of: a) searching and obtaining them from a LOR, b) modifying them and c) offering them back to the LOR is lower than the sum of the costs to develop them (as non-reusable LOs) from the scratch

- **Case 3:** The process of reusing a sequence of LOs (without any modifications) for a new learning activity is cost effective only if the sum of the costs to search and obtain them from a LOR is lower than the sum of the costs to develop them from the scratch as reusable LOs and offer them to the LOR.
- **Case 4:** The process of reusing a sequence of LOs (with modifications) for a new learning activity is cost effective only if the sum of the costs of: a) searching and obtaining them from a LOR and b) modifying them is lower than the sum of the costs to develop them from the scratch as reusable LOs and offer them to the LOR.

For cases 2 and 3, we should mention that Additional Cost for Reusable LO (ADC4RLO) could be reduced to practically zero provided that the particular LO is frequently reused.

An essential cost of the LOs reuse process is the cost of searching and obtaining LOs from LORs. For this purpose, it is important that the LOs process of reuse is supported by effective LORs that can significantly facilitate their end users to narrow their search results and select more easily LOs for reuse within a given learning activity. This will substantially lower the costs for searching and obtaining LOs from the LORs and will make the LOs reusability process more cost effective. Moreover, when modifications to the LOs are needed these increase significantly the cost compared to the cost needed to create the LO from scratch and reduce the potential cost benefits of reuse. Therefore, further analysis would be needed to study under which circumstances LO modifications are costs effective over LO development from scratch. This observation supports the need for LORs to stimulate the versioning and its sharing among LOR users. Finally, possible automatic modifications (i.e., automatic LO modification for different disability categories) can significantly lower the cost of LOs reuse.

Conclusions

The main advantage of Learning Objects in Technology-enhanced Learning has been claimed to be their potential for component-based reuse in different learning settings. Nevertheless, there are only sporadic efforts to study issues related to LOs reuse that would allow interested parties (people, organizations and initiatives) to assess and implement systematic LOs reuse. In this paper, we have studied the concept of LOs reuse within the context of learning activities design and development, we studied and discussed the limitations of existing proposals for LOs reuse and we proposed a thorough workflow for LOs lifecycle that can capture LOs reuse processes. Based on this workflow, we proposed a set of metrics for measuring the cost of LOs reuse as a process rather than measuring only the potential reusability of individual LOs. This is an important issue for large scale deployment of the LO paradigm, since it contributes towards assessing the conditions for LOs reuse being cost effective. The proposed metrics bare the potential for cost benefit analysis of the LOs reuse process from interested parties within the framework of Open Resources Initiatives.

Acknowledgments

The work presented in this paper has been partly supported by the PATHWAY Project that is funded by the European Commission's 7th Framework Programme, Supporting and coordinating actions on innovative methods in science education: teacher training on inquiry based teaching methods on a large scale in Europe (Contract No: SIS-CT-2010-266624).

References

- Bailey, C., Zalfan, M., T., Davis, H., C., Fill, K., & Conole, G. (2006). Panning for Gold: Designing Pedagogically inspired Learning Nuggets, *Educational Technology & Society*, 9(1), 113–122.
- Beetham, H. (2007). An approach to learning activity design. In Beetham, H. & Rhona, S. (Eds.) *Rethinking pedagogy for a digital age: designing and delivering e-learning* (pp. 26–40). London: Routledge.
- Campbell, L. (2003). Engaging with the Learning Object Economy. In Littlejohn, A. (Ed.) *Reusing Online Resources: A Sustainable Approach to eLearning* (pp. 35–45). London: Kogan Page.
- Caswell, T., Henson, S., Jensen, M., & Wiley, D. (2008). Open Educational Resources: Enabling universal education. *The International Review of Research in Open and Distance Learning*, 9(1), 1–11.

- Cervera, J.F., López-López, M.G., Fernández, C., & Sánchez-Alonso, S. (2009). Quality Metrics in Learning Objects, In Sicilia, M.-A. & Lytras, M.D. (Eds.) *Metadata and Semantics* (pp. 135–141). Springer.
- Chidamber, S.R., & Kemerer, C.F. (1994). A metrics suite for object oriented design. *IEEE Transactions on Software Engineering*, 20(6), 476–493.
- Collis, B., & Strijker, A. (2004). Technology and human issues in reusing learning objects. *Journal of Interactive Media in Education*, Retrieved May 1, 2011 from <http://jime.open.ac.uk/2004/4/collis-2004-4.pdf>.
- Colossus: Reuse and Repurposing (2005). *Guide and Case Studies Report, COLOSSUS Project, JISC Exchange for Learning (X4L) Programme*. Retrieved January 12, 2011, from <http://www.strath.ac.uk/projects/colossus/>.
- Conole, G. (2007). Describing learning activities: Tools and resources to guide practice. In H. Beetham & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age* (pp. 81 – 91). Oxford: RoutledgeFalmer.
- Conole, G. & Fill, K. (2005). A learning design toolkit to create pedagogically effective learning activities. *Journal of Interactive Media in Education*, Retrieved May 1, 2011 from <http://jime.open.ac.uk/2005/08/>.
- Cuadrado, J. & Sicilia, M. (2005). Learning objects reusability metrics: Some ideas from software engineering. In Grout, V., Oram, D., & Picking, R., (Eds.), *Proceedings of the International Conference on Internet Technologies and Applications*, Wrexham, UK: North East Wales Institute.
- Currier, S., Barton, J., O’Beirne, R. & Ryan, B. (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. *ALT-J*, 12(1), 5-20.
- Duval, E. & Hodgins, W. (2003). A LOM research agenda. In Hencsey, G., White, B., Robin Chen, Y., Kovács, L. & Lawrence, S. (Eds.). *Proceedings of the Twelfth International Conference on World Wide Web* (pp. 1-9). Budapest, Hungary
- Falconer, I., Conole, G., Jeffery, A. & Douglas, P. (2006). *Learning Activity Reference Model – Pedagogy. LADiE pedagogy guide*, Retrieved May 12, 2011 from http://misc.jisc.ac.uk/refmodels/LADiE/www.elframework.org/refmodels/ladie/guides/LARM_Pedagogy30-03-06.doc.
- Frakes, W. & Terry, C. (1996). Software Reuse and Reusability Metrics and Models, *ACM Computing Survey*, 28(2), 415-435.
- Gehringer, E., L. Ehresman, S. Conger, and P. Wagle. (2007). Reusable Learning Objects Through Peer Review: The Expertiza approach. Retrieved May 12, 2011 from http://innovateonline.info/pdf/vol3_issue5/Reusable_Learning_Objects_Through_Peer_Review-_The_Expertiza_Approach.pdf.
- Hafejh, M., Mili, A., Yacoub, S. & Addy, E. (2002). *Reuse-Based Software Engineering*. Canada: John Wiley.
- IEEE LTSC (2002). *Draft Standard for Learning Object Metadata*. IEEE Learning Technology Standards Committee (LTSC). Retrieved January 12, 2011, from http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf.
- Kay, R. H., & Knaack, L. (2007). Evaluating the learning in learning objects. *Open Learning*, 22 (1), 5-28
- Mat Noor, S. F., Yusof, N., & Mohd Hashim, S. Z. (2009). A Metrics Suite for Measuring Reusability of Learning Objects. *Proceedings of the 9th International Conference on Intelligent Systems Design and Applications* (pp. 961-963). Washington: IEEE Computer Society.
- Mascena, J. C. C. P., Almeida, E. S. & Meira S. R. L. (2005). A Comparative Study on Software Reuse Metrics and Economic Models from a Traceability Perspective. In Zhang, D., Khoshgoftaar, T. & Shyu, M. (Eds.). *Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI)* (pp. 72-77). Nevada, USA.
- McGreal R. (Ed.) (2004). *Online Education Using Learning Objects*. London: Routledge/Falmer.
- Palmer, K. & Richardson, P. (2004). Learning Object Reusability – Motivation, Production and Use. *11th International Conference of the Association for Learning Technology (ALT)*. Devon: University of Exeter.
- Polsani, P. (2003). Use and Abuse of Reusable Learning Objects. *Journal of Digital Information*, 3(4), Retrieved May 1, 2011 from <http://journals.tdl.org/jodi/article/view/89/88>.
- Poulin, J., Caruso, J., & Hancock, D. (1993). The business case for software reuse. *IBM Systems Journal*, 32(4), 567-594.
- Rensing, C., Bergsträßer, S., Hildebrandt, T., Meyer, M., Zimmermann, B., Faatz, A., Lehmann, L. & Steinmetz, R. (2005). Re-Use, Re-Authoring, and Re-Purposing of Learning Resources - Definitions and Examples. *Technical Report KOM-TR-2005-02*. Germany: Darmstadt University of Technology.
- Sampson, D. & Papanikou, C. (2009). A Framework for Learning Objects Reusability within Learning Activities. *Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies (ICALT 2009)* (pp.32-36). Riga, Latvia.
- Van Assche, F. & Vuorikari, R. (2006). A Framework for Quality of Learning Resources. In Ehlers, U. & Pawłowski, J.M. (Eds.) *European Handbook for Quality and Standardization in E-Learning* (pp. 443-456). Berlin: Springer.

Washizaki, H., Yamamoto, Y. & Fukazawa, Y. (2003). A Metrics Suite for Measuring Reusability of Software Components. *Proceedings of the 9th IEEE International Symposium on Software Metrics* (pp. 211-223). Sydney, Australia: IEEE Computer Society.

Weitl, F., Kammerl, R. & Göstl, M. (2004). Context Aware Reuse of Learning Resources. In Cantoni, L. & McLoughlin, C. (Eds.) *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 2119-2126). VA: AACE.

Wiley D. A. (2002). *Instructional Use of Learning Objects*. USA: Agency for Instructional Technology.

Zimmermann, B., Bergsträßer, S., Rensing, C. & Steinmetz, R. (2006). A Requirements Analysis of Adaptations of Re-Usable Content. In Pearson, E. & Bohman, P. (Eds.) *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 2096-2010). Chesapeake, VA: AACE.

Zimmermann, B., Meyer, M., Rensing, C. & Steinmetz, R. (2007). Improving Retrieval of Reusable Learning Resources by Estimating Adaptation Effort backhouse. *First International Workshop on Learning Object Discovery & Exchange (LODE'07)* (pp. 46-53). Crete, Greece.

Jordanian Pre-Service Teachers' and Technology Integration: A Human Resource Development Approach

Jamal Abu Al-Ruz and Samer Khasawneh

The Hashemite University, Faculty of Educational Sciences, Department of Curriculum and Instruction, Zarqa, Jordan 13133 // jamalruz@hu.edu.jo // samer70802@yahoo.com

ABSTRACT

The purpose of this study was to test a model in which technology integration of pre-service teachers was predicted by a number of university-based and school-based factors. Initially, factors affecting technology integration were identified, and a research-based path model was developed to explain causal relationships between these factors. The results supported the hypothesized causal model. The model parameter estimates clearly revealed that a number of factors influenced pre-service teachers' technology integration in their field training. With regard to the university-based factors, the modeling of technology was a highly influential factor impacting pre-service teachers' technology self-efficacy, technology proficiency, and usefulness of technology. Technology self-efficacy was the most important factor with the highest direct effect on technology integration. With regard to the school-based factors, the support structure was the most influential factor with the highest direct effect on technology integration. This study provides some evidence that this model is helpful in determining pre-service teachers' efforts to integrate technology into their classroom practice during field training.

Keywords

Pre-service teachers, Technology integration, Human resource development, Jordan

Introduction

In recent years, the use of computer technology in education has gained global acceptance. Computer technology is widely used as an instructional tool in almost every teaching-learning setting and its use is continuing to expand (Hogarty, Lang, & Kromrey, 2003). There is a general belief that technology integration in the curriculum may result in improvement of classroom instruction (Libscomb & Doppem, 2004) and ultimately may provide students with the needed skills to survive and compete in the twenty-first century digital society (Norris, Sullivan, Poirot, & Soloway, 2003). Further, it may improve students' learning (Mills & Tincher, 2003); critical thinking skills (Harris, 2002); and achievement, motivation, and attitudes (Waxman, Lin, & Michko, 2003).

Technology can provide powerful tools for students' learning, but its value depends upon how effectively school teachers use it to support instruction in the classroom (Fulton, Glenn, & Valdez, 2004). One promising area of research involves the study of technology integration in the classroom by pre-service teachers. Pre-service teachers are viewed as the transmitters of up-to-date knowledge and can effectively link theory into practice. The ability of pre-service teachers to integrate technology into the curriculum is needed to guarantee their future success and the success of their students. To this end, many teacher-education programs are concerned with how to properly provide pre-service teachers with the technology-related attitudes and skills needed to integrate technology into classroom practices (Wilson, 2003). It is well documented in the literature that the teacher-education courses that expose pre-service teachers to technology play a major role in pre-service teachers' overall use of technology and may assist them in learning to integrate technology into their future classroom practice (Collier, Weinburgh, & Rivera, 2004; Pope, Hare, & Howard, 2002).

Models of technology use by pre-service teachers have been developed over the past few years. For example, Venkatesh, Morris, Davis, and Davis (2003) developed the Unified Theory of Acceptance and Use of Technology (UTAUT). They suggested that eight elements play an important role in technology acceptance: gender, age, experience, voluntariness of use, performance expectancy, effort expectancy, social influence, and facilitating conditions. These variables were found to predict 70% of the variance in user intentions toward computer technology. Yuen and Ma (2002) used the Technology Acceptance Model (TAM) with pre-service teachers to examine the influences of perceived usefulness and perceived ease of use on their intention to use. The results of the study indicated that perceived usefulness had a significant positive effect on intention and usage but not on perceived ease of use. Likewise, Smarkola (2007) and Ma, Andersson, and Streith (2005) used a modified version of the TAM

to examine determinants for pre-service teachers' use of computer technology. They discovered that perceived usefulness and ease of use of the technology were the key factors.

Dexter and Riedel (2003) found that pre-service teachers' comfort with technical skills (e.g., word processors and Internet browsers) and availability of computers at school site was rated the highest in their effect on technology integration in the teaching process. In their study, based on individual interviews, pre-service teachers indicated that technology was not modeled by instructors via the courses taught. Chen (2004) found that pre-service teachers increased their confidence in using computer technology by having experiences from a previous computer course. Chen (2004) asserts that "teachers need to have the confidence and positive attitudes towards computers that will motivate them to integrate computers into their instructional strategies" (p. 50). Moreover, Anderson and Maninger (2007) found that pre-service teachers' self-efficacy was the best predictor of technology use in the classroom. Further, Smarkola (2008) used the decomposed theory of planned behavior to examine pre-service teachers' intentions to use computer technology in their teaching. The results indicated that usefulness of computers and computer confidence were the best predictors.

Based on the above discussion, it is obvious that regardless of the level of available infrastructure and support from administration, there is concern as to whether pre-service teachers are prepared to integrate the technology that is available to them into teaching (Brown & Warschauer, 2006; Firek, 2002; Ma, Andersson, & Streith, 2005). Some of these factors were modeling of technology, computer self-efficacy, computer proficiency, and perceived usefulness of technology. The present research is an attempt to modify previous research models by proposing a new model that is more relevant to pre-service teachers.

Teacher-education research suggests that pre-service teachers need to observe university faculties modeling technology in their courses to learn how technology can be effectively used to enhance instruction (O'Bannon & Judge, 2004; Schrum, Skeele, & Grant, 2003). It is through these courses that pre-service teachers are to receive all the training that they will need to integrate technology into their teaching upon entering their classroom practice (Banister & Vannatta, 2006). On the other hand, it is suggested that modeling technology in university courses may improve students' technology self-efficacy, technology proficiency, and the perceived usefulness of technology. Further, overall support and technology availability also influenced the use of technology in the classroom.

The first university-based factor is technology self-efficacy. Technology self-efficacy refers to pre-service teachers' perceptions of their ability to use technology effectively in the classroom. According to Social Learning Theory (Bandura, 1997), successful past experience with technology (vicarious experiences) would be expected to lead to higher self-efficacy whereas poor past performance would tend to lower self-efficacy (Wood & Bandura, 1989). It is therefore likely that those pre-service teachers who receive some type of training about how to use relevant dimensions of technology may develop and report more positive efficacy beliefs (self-confidence levels) than those who do not receive such training (Dawson & Rakes, 2003). These high self-efficacy beliefs, in turn, would lead to pre-service teachers' successful integration of technology in instruction (Bandura, Adams, & Beyer, 1977; Wall, 2004).

The second university-based factor is technology proficiency. Pre-service teachers are expected to be knowledgeable about current technology and how it can be used to promote learning (Jacobsen, Clifford, & Friesen, 2002). It is well established in the literature that improvements in university students' technology proficiencies were reported during their courses of study in which technology was integrated (Anderson & Boarthwick, 2002). This technology proficiency, in turn, is one of the most important characteristics influencing pre-service teachers' success at integrating technology in instruction (Hernandez-Ramos, 2005; Kanaya, Light, & Culp, 2005).

The third university-based factor is the perceived usefulness of technology, meaning the degree by which pre-service teachers feel technology is useful for present and/or future work. When pre-service teachers are exposed to technology during coursework, various aspects of such technology may enhance their perceptions of the usefulness of the technology in their future jobs. In turn, such perceived usefulness of technology plays a critical role in predicting integration of technology in the classroom (Mathieson, Peacock, & Chin, 2001).

Previous research has also emphasized the importance of other factors impacting technology integration in the classroom. Among these highly influential factors is overall support and technology availability. Overall support is support that is technical in nature or comes from teachers and administrators. Both types of support have often been

considered to be influential factors in teachers' technology-integration practices (Grant, Ross, Wang, & Potter, 2005; O'Dwyer, Russell, & Bebel, 2004). Moreover, research has also emphasized the importance of the support that comes from the school principal as well as from cooperating teachers, which is the primary stimulus for incorporating technology in the classroom (Zhao & Frank, 2003). The other factor that is important to pre-service teachers' technology integration is the availability of the technology. Access to technology has been thought of as a main obstacle for technology integration for the last decade (Culp, Honey, & Mandinach, 2003). Without adequate technology, pre-service teachers may have little opportunity to integrate technology into the classroom (Morris, 2002). Filed placements in technology-enriched environments have been found in the research to be a positive factor contributing to technology integration (Karagiorgi, 2005). Although pre-service teachers are prepared to use technology in their field placements, the lack of support and technology availability (access to computers and software) also play an important role in integrating technology into the classroom (Vannatta & Fordham, 2004; Wozney, Venkatesh, & Abrami, 2006). In short, pre-service teachers' integration of technology is indeed influenced by factors found in both the university environment and field-training environment.

Statement of the problem

How to prepare pre-service teachers to integrate technology in the classroom has been a subject of concern in recent years. Even though the integration of technology in the classroom has the potential to enhance students' learning, the research continues to report that pre-service teachers are not utilizing technology in the classroom during their field training (Morris, 2002). The primary purpose of this study is to develop and test two path models, one related to the university environment, and the other related to the school environment. The first path model describes an antecedent variable (the modeling of technology in teacher-education courses) that influence multiple university-based variables (technology self-efficacy, technology proficiency, and perceived usefulness of technology) of pre-service teachers' efforts to integrate technology in the classroom. The second path model tests the effects of two school-based variables (overall support and technology availability) on pre-service teachers' effort to integrate technology in the classroom.

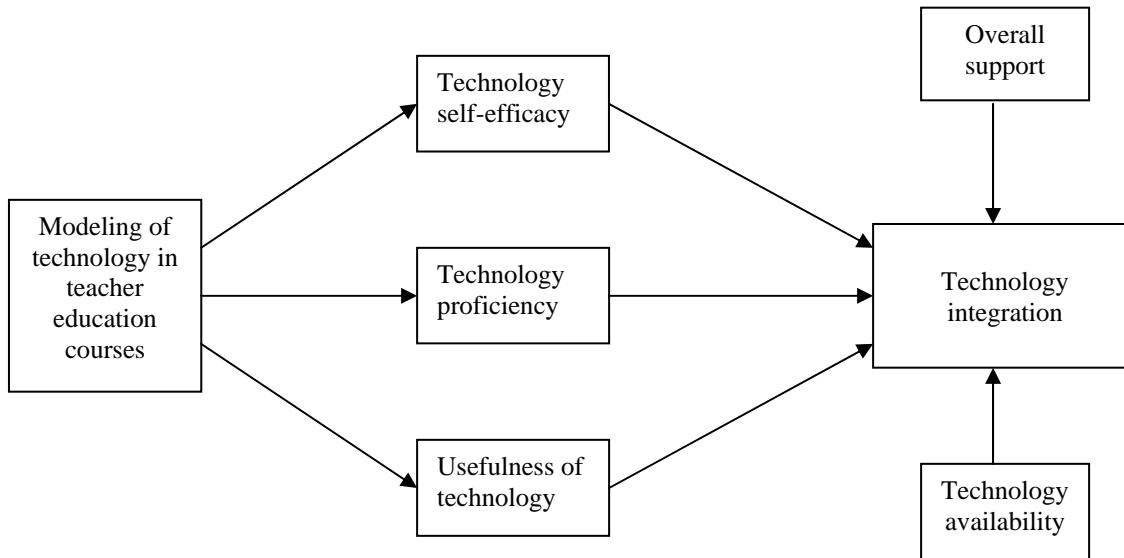


Figure 1. A model of pre-service teachers' integration of technology

Overview of the model

The first research model hypothesized a positive link from the modeling of technology in teacher-education courses to three university-based factors (technology self-efficacy, technology proficiency, and usefulness of technology). The aforementioned university factors are also hypothesized to positively influence technology integration. The second research model hypothesized a positive link from two school-based factors (overall support and technology availability) to technology integration. The two path models are isolated but are related through their impact on technology integration, meaning that in order for the first path model to take place, the second path model should be in place. Figure 1 presents the hypothesized relationships.

Methodology

Research design

The design employed in this study was a descriptive survey research design in which factors impacting pre-service teachers' technology integration were investigated through survey instrument. Path models of the direct effects of prediction variables were tested. To assess the adequacy of the models' fit, path analysis was conducted using LISREL 8.51 procedures (Joreskog & Sorbom, 1993).

Study context

The present study took place in one of Jordanian public universities. The undergraduate program prepares classroom teachers in capstone courses related to the curriculum, teaching sources, teaching methods, and technology use in a variety of subjects including Arabic language, Islamic studies, social studies, science, math, and vocational education. Within this program, faculty members utilize instructional technology (e.g., Blackboard and Integrity systems) and micro-teaching to deliver high-quality instruction. Moreover, students are required to interact with this technology in the form of discussion boards, digital drop boxes, video watching, and presentations. On the last semester prior to graduation, pre-service teachers are required to attend practicum training in hosting schools, teaching actual classes. Pre-service teachers usually teach all courses for grades 1 to 4 five days a week from 7:30 AM to 2:15 PM. Further, pre-service teachers are required to spend three hours per week on the university campus, meeting with faculty members to discuss their field experience and effective methods of integrating technology into classroom instruction.

Population and sample

The target population for this study was the 1120 classroom teachers who attended the teacher-education program at a public university in Jordan for the academic years 2006 through 2009. The sample of this study comprised 1,008 pre-service teachers who volunteered to participate in the study. All pre-service teachers were seniors in their final semester prior to graduation. The study sample was mostly females (83%).

Instrumentation

A 36-item survey was used in this study. The instrument was developed from several sources. The first part of the instrument, the technology self-efficacy sub-scale with 10 items, was adopted from the computer attitude measure developed and tested by Gressard and Lioyd (1986). An example of items includes "I have a lot of self-confidence when it comes to working with computer technology" and "I think using a computer would be very easy for me." Few of the scale items were reworded (e.g., the term "very difficult" was changed to "very easy"). This self-efficacy subscale has been found to be a valid and reliable measure. The coefficient alpha reliability for the subscale was .91 (Gressard & Lioyd, 1985) and .89 (Gressard & Lioyd, 1986). Respondents were asked to rate each item along a five-point Likert scale as follow: (1) strongly disagree, (2) disagree, (3) uncertain, (4) agree, and (5) strongly agree. A higher score on the total score indicates a more positive self-efficacy beliefs related to computer technology. The second part of the instrument, the perceived usefulness of computer technology, is adapted from the TAM (Davis, Bagozzi, & Warshaw, 1989) and included six items. This subscale is intended to assess the learners' perceptions of

the usefulness of technology. An example of items in this subscale includes “using technology would enhance my effectiveness on the job” and “using computers would make it easier to do my job.” The coefficient alpha reliability for the subscale was .94 (Arbaugh, 2000), emphasizing its reliability for future use. Respondents were asked to rate each item along a five-point Likert scale as follow: (1) strongly disagree, (2) disagree, (3) uncertain, (4) agree, and (5) strongly agree. A higher score on the total score indicates a more positive perceived usefulness of technology.

An Arabic version of the two subscales was achieved through a standard three-step protocol. First, the two subscales were translated from English into Arabic language by a professional scholar fluent in both English and Arabic. Second, the two subscales were translated back from Arabic into English language by a second professional scholar, also competent in both English and Arabic. In the final step, a third professional scholar, fluent in both English and Arabic, compared and evaluated the original English and back-translated copies in order to verify the accuracy and validity of translation. Then, nine specialists in education and technology reviewed the two developed subscales, and two of them asked for minor modifications. The final copy of the questionnaire took these modifications into consideration.

The researchers developed the other five scales used in the study with the assistance of several content judges who had expertise in the use of technology in the classroom. Scale items were drafted by the researchers and submitted to several content judges for review. Based on their feedback, items were added, dropped, or reworded where necessary. A pilot-study preliminary questionnaire with a group of 25 students and instructors was conducted. Pilot testers read the items aloud in order for the researchers to determine if their interpretations of items matched the intended meanings. Feedback from this pilot study led to minor modifications in the wording of several items. The five subscales along with examples of items were as follow: modeling of technology (technology integration was discussed in one or more of my courses this semester); technology proficiency (I have the skills necessary to use computers for instruction); technology integration (I had opportunities to integrate technology into my student teaching experiences); overall support (the classroom teacher I worked with was supportive of using computers in the classroom); and technology availability (materials such as software and printer supplies) for classroom (computers are readily available for use).

Exploratory factor analysis was conducted to provide some evidence of construct validity for the measures. In the present study, exploratory common factor analysis was used to identify the underlying latent structure of the data. The criteria for determining how many factors to extract included the eigenvalue greater than one rule and a visual inspection of the scree plot. The initial analysis was run without specifying how many factors to retain. This procedure resulted in seven factors with 36 items, explaining 54.34% of the common variance. Items were retained on factors with a minimum loading of .30, but were not retained if they had a cross loading above .20. Factor loadings for items retained in this solution ranged from .33 to .83, with an average loading of .61 on major factor and .05 on the rest of the factors. The results of the factor analysis closely paralleled the hypothesized variables and the following scales and items emerged: modeling of technology (6 items) technology self-efficacy (7 items), technology proficiency (5 items), usefulness of technology (6 items), technology integration (4 items), overall support (4 items), and technology availability (4 items). All of these scales used a five-point Likert scale, with values ranging from 1 (strongly disagree) to 5 (strongly agree). Estimates of reliability using Cronbach's alpha were acceptable for all scales ranging from .73 to .90 (see Tables 1 & 2).

Data collection

The researchers, who have been major participants in the teacher-education program for the past five years, contacted pre-service teachers in person at school and through on-campus meetings to participate in the study. The researchers explained to the participants the purpose of the study and encouraged them to read the statements carefully before ticking the appropriate choice. The volunteer participants were also insured confidentiality and anonymity. Further, participants were also informed that the instrument takes approximately 15 to 20 minutes to complete. Finally, instruments were handed out and collected by the researcher.

Data analysis

The Pearson product moment correlation coefficient was the statistical measure used to determine the strength of the associations among the hypothesized variables (Table 1). An alpha level of .05 was used to determine the

significance of relationships. All variables were tested using covariance matrices generated by PRELIS and utilized the maximum-likelihood method to estimate parameters in the path model. In path models there are two types of variables: exogenous and endogenous (Klem, 1995). Exogenous variables have no causal links toward them from other variables in the model (e.g., modeling of technology, overall support, and technology availability). Additionally, the value of the exogenous variable is not explained by those other variables. In contrast, an endogenous variable has causal links coming toward it in the model (e.g., technology integration, technology self-efficacy, technology proficiency, and usefulness of technology), and its value is explained by one or more of the other variables (Schumacker & Lomax, 2004). Also, endogenous variables can be both dependent and independent variables (e.g., technology self-efficacy) (Klem, 1995).

Results

Correlation analysis

Prior to structural modeling (path analysis), the correlations among the variables were obtained. The correlation matrix shown in Table 1 indicated that the modeling of technology was positively associated with technology self-efficacy ($r = .50, p < .01$); technology proficiency ($r = .49, p < .01$), and usefulness of technology ($r = .52, p < .01$). Technology self-efficacy was positively associated with technology integration ($r = .52, p < .01$); technology proficiency was positively associated with technology integration ($r = .39, p < .01$); and usefulness of technology was positively associated with technology integration ($r = .36, p < .01$). As shown in Table 2, technology availability was positively associated with technology integration ($r = .39, p < .01$), and overall support was positively associated with technology integration ($r = .44, p < .01$).

Table 1. Reliabilities, means, standard deviations, and correlations for university variables

| Scale | α | Mean | SD | N | 1 | 2 | 3 | 4 | 5 |
|--------------------------|----------|------|-----|------|-------|-------|-------|-------|---|
| 1 Modeling of technology | .89 | 3.39 | .80 | 995 | — | | | | |
| 2 Self-efficacy | .90 | 3.77 | .64 | 999 | .50** | — | | | |
| 3 Proficiency | .82 | 3.70 | .66 | 992 | .49** | .52** | — | | |
| 4 Usefulness | .86 | 3.19 | .78 | 974 | .52** | .37** | .45** | — | |
| 5 Technology integration | .80 | 4.07 | .59 | 1002 | .38** | .52** | .39** | .36** | — |

** Correlation is significant at the 0.01 level (two-tailed).

Table 2. Reliabilities, means, standard deviations, correlations for school variables

| Scale | α | Mean | SD | N | 1 | 2 | 3 |
|---------------------------|----------|------|-----|------|-------|-------|---|
| 1 Technology availability | .87 | 3.58 | .72 | 1002 | — | | |
| 2 Overall support | .73 | 3.59 | .71 | 1000 | .24** | — | |
| 3 Technology integration | .80 | 4.07 | .59 | 1002 | .39** | .44** | — |

** Correlation is significant at the 0.01 level (two-tailed).

Path analysis

Six fit indices were examined in this study, including the chi-square test. These indices were the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), the comparative fit index (CFI), the non-normed fit index (NNFI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR) (Byrne, 1998). A value of .90 or above for the GFI and AGFI is usually recommended for an acceptable level of fit (Hair et al., 1998). Finally, RMSEA values less than .10 represent models with a good fit to the data (Byrne, 1998). Similar to the RMSEA, the SRMR represents the square root of the mean residuals between the implied model and the data. Values less than .05 are generally indicative of a good fit of the model to the data (Byrne, 1998). The last two fit indices (CFI and NNFI) are considered incremental fit indices because they measure the proportionate improvement in fit of the proposed model relative to a baseline represented by the null model. These measures have the advantage of being less influenced by sample size when compared to other indices, such as GFI. Generally, values above .90 are considered sufficient (Byrne, 1998).

The chi-square value (χ^2 (339) = 1621.07, $p < .01$) was significant. A significant chi-square value indicates that the proposed path model does not completely fit the observed covariances and correlations (Hair et al., 1998). However, the chi-square by itself should not be used as the sole indicator of model fit due to its high sensitivity to sample size and violation of multivariate normality. Therefore, consideration of other fit indices is considered essential. For example, the values for GFI (.91), AGFI (.90), CFI (.92), and NNFI (.91) indicated that the model fit the data sufficiently (Byrne, 1998). The RMSEA (.06) and the SRMR (.04) value indicated that there was a minimal amount of error associated with the tested path model (Byrne, 1998). The standard errors of all the estimates were small enough to say that the estimates are relatively precise. The t -values for the paths above were the absolute value of 1.96, indicating that paths were significant at the .01 level (Joreskog & Sorbom, 1989). Finally, the modification indices provided by LISREL did not suggest any significant changes to improve the model, implying that this model fits the data relatively well. Eight separate paths were tested in this model. The results of the path analysis are summarized in Figure 2, which displays the standardized path coefficients (beta weights) as well as the explained variance (R^2) for the dependent variables. As can be seen, all eight of the hypothesized paths were supported ($p < .01$).

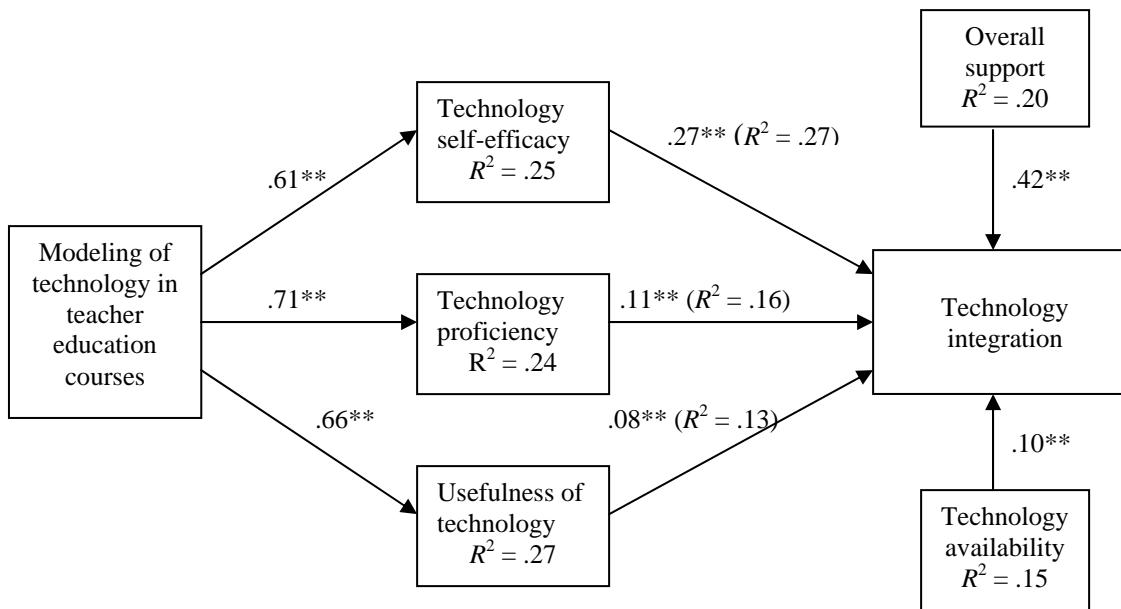


Figure 2. A model of pre-service teachers' integration of technology

With regard to the university-based factors, the model shows that the modeling of technology has a direct positive effect on technology self-efficacy (beta = .61), technology proficiency (beta = .71), and usefulness of technology (beta = .66). Technology self-efficacy has a direct positive effect on technology integration (beta = .27); technology proficiency has a direct positive effect on technology integration (beta = .11); and usefulness of technology has a direct positive effect on technology integration (beta = .08). With regard to the school-based factors, technology availability has a positive direct effect on technology integration (beta = .10) and overall support has a positive direct effect on technology integration (beta = .42) (see Figure 2). Overall, this model had an adequate predictive power as shown by the R^2 statistic. From the first path model, modeling of technology explained 25% of the variance in technology self-efficacy, 24% of the variance in technology proficiency, and 27% of the variance in usefulness of technology. Furthermore, 27% of the variance in technology integration was explained by technology self-efficacy, 16% of the variance was explained by technology proficiency, and 13% of the variance was explained by usefulness of technology. From the second path model, technology availability explained 15% of the variance in technology integration, and overall support explained 20% of the variance in technology integration.

Discussion

This study represents a research-based effort to evaluate one critical factor (the modeling of technology), leading to the development of several consequence factors (technology self-efficacy, technology proficiency, and usefulness of technology), which are fundamental antecedents to pre-service teachers' technology integration efforts. The theory and research suggest that technology integration is also influenced by situational factors in the field-training practice, including technology availability and overall support to use technology. The present study developed two research-based path models, hypothesizing a direct positive link from faculty modeling to three university-based factors (technology self-efficacy, technology proficiency, and usefulness of technology) that are mediators to technology integration. Also, the study hypothesized a direct positive link from two school-based factors (overall support and technology availability) to technology integration.

The results are consistent with the conceptualization of technology self-efficacy, technology proficiency, and usefulness of technology as mediators between the modeling of technology and technology integration by pre-service teachers. Specifically, the modeling of technology was associated with higher levels of technology self-efficacy, higher levels of technology proficiency, and higher levels of perceived usefulness of technology. Technology self-efficacy, technology proficiency, and usefulness of technology, in turn, were positively associated with pre-service teachers' technology integration efforts. The results are consistent with previous research that pre-service teachers need to observe university faculties modeling technology in their courses to learn how technology can be effectively used to enhance instruction (Banister & Vannatta, 2006).

In this study, university faculties modeled technology in instruction (e.g., used the Blackboard system), which in turn, impacted students' confidence in interacting with the technology via the discussion board, digital drop box, and e-mails. This finding is congruent with a social-learning perspective on the development and role of self-efficacy as contributor to the direction, intensity, and persistence of effort related to the use of technology in the classroom. (Bandura, Adams, & Beyer, 1977). Moreover, the modeling of technology impacted students' proficiency and skills in dealing with various technological tools during instruction. It is well established in the literature that improvements in university students' technology proficiencies were reported during their courses of study in which technology was integrated (Topper, 2004). This technology proficiency, in turn, is one of the most important characteristics influencing pre-service teachers' success at integrating technology in their instruction (Hernandez-Ramos, 2005). Further, the modeling of technology by faculty members affected students' perceived usefulness of technology in their present role as students and in their future careers as classroom teachers (Dawson & Rakes, 2003).

Overall, these results suggest that the higher the modeling of technology in the university teacher-education courses, the higher the mediating factors (technology self-efficacy, technology proficiency, and perceived usefulness of technology) and the higher the mediating factors the higher pre-service teachers are integrating technology in their field training classrooms.

The second path model speculated that technology integration is influenced by two school-based factors (overall support and technology availability). Results of the study supported the hypothesized model. In this study, pre-service teachers received support from technicians, teachers, and principals, which, in turn, affected their technology integration in instruction. These results are consistent with previous research indicating that such support is often considered to be an important factor in teachers' technology integration practices (Zhao & Frank, 2003). The other factor that is important to pre-service teachers' technology integration efforts is the availability of technology. In this study, pre-service teachers indicated that technology was available in the practising schools (e.g., computers, printers, software), which had an impact on the teachers' ability to integrate technology in instruction. This finding is consistent with previous research emphasizing that without adequate technology, pre-service teachers have little opportunity to integrate technology into the classroom (Morris, 2002). Based on these results, we can speculate that the higher the support structure and technology availability, the higher the technology integration efforts by pre-service teachers.

Conclusions and recommendations

In conclusion, the importance of the present study lies essentially in gaining a deeper understanding of the factors that influence pre-service teachers' effort to integrate technology in the field-training classrooms, which can help

administrators in higher education settings recognize the importance of faculty usage of technology in university courses to foster positive technology self-efficacy beliefs, proficiency, and usefulness. Further, this study informs professionals in K–12 schools of the status of their current support structure and the availability of technology to ensure successful integration of technology by pre-service teachers. Finally, we suggest a number of practical and theoretical recommendations for the field of study.

From the practical standpoint, faculty members in higher education institutions should pay close attention to setting conditions that enhance the development of pre-service teachers' technology self-efficacy, technology proficiency, and usefulness of technology. This includes the modeling of technology in the courses that they teach. Thus, preparatory activities such as familiarizing students with the technology, discussing how it will be used to meet learning objectives, and providing opportunities to experience some early successes with the technology appear to be important strategies contributing to the formation of these factors and motivating pre-service teachers to integrate technology into their field training. Another recommendation is that university administration set policies demanding faculty members and pre-service teachers to attend technology training programs (e.g., ICDL, IC3) to enhance the teaching-learning process, have technology mentors on-campus to better meet the needs and questions of pre-service teachers as they progress through their field training, and ensure that pre-service teachers attend field-training in schools that support the use of technology in instruction and have adequate technology available on-site. The final recommendation is that the Ministry of Education ensure that their schools are prepared with the technology needed to deliver effective instruction (e.g., adequate computers and Internet connections) and that school teachers and administrators support the use of technology in instruction.

From a theoretical standpoint, researchers should attempt to fully develop a path model of structural relations between constructs investigated in this study. This can be done through interviews and focus groups that include faculties, pre-service teachers, cooperating teachers, and school principals to determine other factors that may contribute to technology integration and other contexts such as individual variables that may play a part in this nomological network. This research can be replicated with all public and private universities in Jordan to confirm the findings in this study. Furthermore, researchers should attempt to test competing models to technology integration practices by pre-service teachers in Jordan to develop theories related to the field of study. Also, national and international researchers should cooperate to study how culture can play a part in these models.

References

- Anderson, C. L., & Boarhwick, A. (2002). Results of separate and integrated technology instruction in preservice training. *Paper presented at the National Educational Computing Conference*, San Antonio, TX.
- Anderson, S. E. & Maninger, R. M. (2007). Pre-service teachers' abilities, beliefs, and intentions regarding technology integration. *Journal of Educational Computing Research*, 37(2), 151–172.
- Bandura, A. (1997). *Self-efficacy: The Exercise of control*. New York: Freeman.
- Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavior change. *Journal of Personality and Social Psychology*, 35(3), 125–139.
- Banister, S., & Vannatta, R. (2006). Beginning with a baseline: Insuring productive technology integration in teacher education. *Journal of Technology and Teacher Education*, 14(1), 209–235.
- Brown & Warschauer (2006). From the university to the elementary classroom: Students' experiences in learning to integrate technology in instruction. *Journal of Technology and Teacher Education*, 14(3), 599–621.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Chen, L. L. (2004). Pedagogical strategies to increase pre-service teachers' confidence in computer learning. *Educational Technology and Society*, 7(3), 50–60.
- Collier, S., Weinburgh, M. H., & Rivera, M. (2004). Infusing technology skills into a teacher education program: Change in students' knowledge about and use of technology. *Journal of Technology and Teacher Education*, 12(3), 447–468.
- Culp, K. M., Honey, M., & Mandinach, E. (2003). *A retrospective on twenty years of education technology policy*. Washington, DC: US Department of Education, Office of Educational Technology.

- Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982–1003.
- Dawson, C., & Rakes, G. C. (2003). The influence of principals' technology training on integration of technology into schools. *Journal of Research on Technology in Education*, 36(1), 29–49.
- Dexter, S., & Riedel, E. (2003). Why improving pre-service teacher educational technology preparation must go beyond the college's walls. *Journal of Teacher Education*, 54, 340–346.
- Firek, H. (2002). One order of educational technology coming up...you want fries with that? *Phi Delta Kappan*, 84(8), 596–597.
- Fulton, K., Glenn, A. D., & Valdez, G. (2004). *Teacher education and technology planning guide*. North Central Regional Educational Laboratory, Learning Point Associates.
- Gressard, C. P., & Loyd, B. H. (1985). Age and staff development experience with computers as factors affecting teacher attitudes toward computers. *School Science and Mathematics*, 85(3), 203–209.
- Gressard, C. P., & Loyd, B. H. (1986). Validation studies of a new computer attitude scale. *Association for Educational Data System Journal*, 19, 295–301.
- Grant, M. M., Ross, S. M., Wang, W., & Potter, A. (2005). Computers on wheels: An alternative to "each one has one." *British Journal of Educational Technology*, 36(6), 1017–1034.
- Hair, J. E., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th Ed.). Upper Saddle River, NJ: Prentice-Hall.
- Harris, C. M. (2002). Is multimedia-based instruction Hawthorne revisited? Is difference the difference? *Education*, 122(4), 839–843.
- Hernandez-Ramos, P. (2005). If not here, where? Understanding teachers' use of technology in Silicon Valley schools. *Journal of Research on Technology in Education*, 38(1), 39–64.
- Hogarty, K. Y., Lang, T. R., & Kromrey, J. D. (2003). Another look at technology use in classrooms: The development and validation of an instrument to measure teachers' perceptions. *Educational and Psychological Measurement*, 63(1), 139–162.
- Jacobsen, M., Clifford, P., & Friesen, S. (2002). Preparing teachers for technology integration: Creating a culture of inquiry in the context of use. *Contemporary Issues in Technology and Teacher Education*, 2(3), 363–388.
- Joreskog, K. G., & Sorbom, D. (1989). *LISREL 8: User's Reference Guide*. Chicago: Scientific Software, Inc.
- Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: User's Reference Guide*. Chicago: Scientific Software, Inc.
- Kanaya, T., Light, D., & Culp, K. M. (2005). Factors influencing outcomes from a technology-focused professional development program. *Journal of Research on Teaching in Education*, 37(2), 313–329.
- Karagiorgi, Y. (2005). Throwing light into the black box of implication: ICT in Cyprus elementary schools. *Educational Media International*, 42(1), 19–32.
- Klem, L. (1995). Path analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.
- Libscomb, G. B., & Doppen, F. H. (2004). Climbing the stairs: Pre-service social studies teachers' perceptions of technology integration. *International Journal of Social Education*, 19(2), 70–87.
- Ma, W.W., Andersson, R. & Streith, K.O. (2005). Examining user acceptance of computer technology: An empirical study of student teachers. *Journal of Computer Assisted Learning*, 21(6), 387–395.
- Mathieson, K., Peacock, E., & Chinn, W. (2001). Extending the Technology Acceptance Model: The influence of perceived user resources. *The Database for Advances in Information Systems*, 32(3), 86–112.
- Mills, S. C., & Tincher, R. C. (2003). Be the technology: A developmental model for evaluating technology integration. *Journal of Research on Technology in Education*, 35(3), 1–20.
- Morris, M. (2002). How new teachers use technology in the classroom. *Paper presented at the Annual Summer Conference of the Association of Teacher Educators*, Williamsburg, VA.
- Norris, C., Sullivan, T., Poirot, J., & Soloway, E. (2003). No access, no use, no impact: Snapshot surveys of educational technology in k-12. *Journal of Research on Technology in Education*, 36(1), 15–27.
- O'Bannon, B., & Judge, S. (2004). Impacting partnership across the curriculum with technology. *Journal of Research on Technology in Education*, 37(2), 198–211.

- O'Dwyer, L., Russell, M., & Bebel, D. (2004). *Elementary teachers' use of technology: Characteristics of teachers, schools, and districts associated with technology use*. Boston, MA: Technology and Assessment Study Collaborative, Boston College.
- Pope, M., Hare, D., & Howard, E. (2002). Enhancing technology use in student teaching: A case study. *Journal of Technology and Teacher Education*, 13(4), 573–618.
- Schrum, L., Skeele, R., & Grant, M. (2003). One college of education's effort to infuse technology: A systematic approach to revisioning teaching and learning. *Journal of Research on Technology in Education*, 35(2), 226–271.
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Smarkola, C. (2007). Technology acceptance predictors among student teachers and experienced classroom teachers. *Journal of Educational Computing Research*, 37(1), 65–82.
- Smarkola, C. (2008). Efficacy of a planned behaviour model: Beliefs that contribute to computer usage intentions of student teachers and experienced teachers. *Computers in Human Behavior*, 24(3), 1196–1215.
- Topper, A. (2004). How are we doing? Using self-assessment to measure changing teacher technology literacy within a graduate educational technology program. *Journal of Technology and Teacher Education*, 12(3), 303–317.
- Vannatta, R. A., & Fordham, N. (2004). Teacher dispositions as predictors of classroom technology use. *Journal of Research on Technology on Education*, 36(3), 253–271.
- Venkatesh, V., Morris, F., Davis, A., & Davis, A. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27, 425–478.
- Wall, A. (2004). *An evaluation of the computer self-efficacy of preservice teachers*. Unpublished doctoral dissertation, Tennessee State University, Nashville, Tennessee.
- Waxman, H. C., Lin, M., & Michko, G. M. (2003). *A model analysis of the effectiveness of teaching and learning with technology on student outcomes*. Retrieved May 1, 2011 from: <http://www.ncrel.org>.
- Wilson, E. (2003). Pre-service secondary social studies teachers and technology integration: What do they think and do in their field experiences? *Journal of Computing in Teacher Education*, 20(1), 29–39.
- Wood, R., & Bandura, A. (1989). Impacts of conception of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology*, 56, 407–415.
- Wozney, L., Venkatesh, V., & Abrami, P. (2006). Implementing computing technologies: Teachers' perceptions and practices. *Journal of Technology and Teacher Education*, 14(1), 173–207.
- Yuen, A.H.K and Ma, W.W.K (2002). Gender differences in teacher computer acceptance. *Journal of Technology and Teacher Education*, 10(3), 365–382.
- Zhao, Y., & Frank, K. A. (2003). Factors affecting technology uses in schools: An ecological perspective. *American Educational Research Journal*, 40(4), 807–840.

An Individualized e-Reading System Developed Based on Multi-representations Approach

Chien-Chuan Ko¹, Chun-Han Chiang², Yun-Lung Lin³ and Ming-Chung Chen²

¹Department of Computer Science and Information Engineering, National Chiayi University, Chiayi City 60004, Taiwan // ²Department of Special Education, National Chiayi University, Chiayi 60004, Taiwan // ³Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taipei, Taiwan // kocc@mail.ncyu.edu.tw // s0940606@mail.ncyu.edu.tw // harrison@ice.ntnu.edu.tw // mtchen@mail.ncyu.edu.tw

ABSTRACT

Students with disabilities encounter many difficulties in learning activities, especially in reading. To help them participate in reading activities more effectively, this study proposed an integrated reading support system based on the principle of multiple representations. An integrated e-reading support system provides physical, sensory, and cognitive support to learners. This system also served as the convenient interface for material developers and instructors. The results of usability evaluation also demonstrated the interface of learner to be friendly and efficient. Thirty fifth- and sixth-grade students with learning disabilities in Taiwan participated in the experiment to explore if this e-reading system could help them to understand natural-science articles. All students read the articles on the e-reading system, with and without cognitive support. The results indicated better comprehension performance when the participants read with cognitive support.

Keywords

Reading difficulties, Cognitive support, Multiple representations, Individualized reading system, Students with disabilities

Introduction

Reading skill is essential to a successful learning activity. However, without any help, it is difficult for learners with special needs to read effectively due to their limitations or disabilities, such as dyslexia, intellectual disability, visual impairment, visual perception difficulties, or palsy. Such disabilities prevent these students from meeting the challenge of the general curriculum (Bender, 2004; Lerner, 2006; Mastropieri & Scruggs, 2002). Therefore, remedying the deficiencies of reading and improving the reading skills, including word recognition and reading comprehension, has been one of the major services provided by schools. Recent studies have supported various effective strategies, such as word-recognition instruction methods (Browder, Spooner, Wakeman, Trella, & Baker, 2006; Hung & Huang, 2006; Van der Bijl, Alant, & Lloyd, 2006), reciprocal teaching (Ledere, 2000; Palinscar & Brown, 1984), and drawing concept maps (Chang, Sung, & Chen, 2002; Guastello, Beasley, & Sinatra, 2000). Though effective, these strategies, however, remain inadequate for certain disabled users. Therefore, researchers explored alternative solutions that would allow users to bypass their disabilities or augment their residual capabilities (Lewis, 1993).

To help special needs students access the reading material, typical assistive technology devices are provided, for example, e-readers, screen-magnifying software, and adaptive computer devices. Popular e-readers, e.g. WYNN Literacy Software Solution and Kurzweil 3000, could read the text for the poor readers. Screen-magnifying software can help people with low vision read electronic texts by enlarging the characters. A switch equipped with a scanning program allows people with severe motor impairment to read the electronic texts (Alliance for Technology Access, 2004). However effective these methods may be, they are of little use for the cognitively challenged, especially for those lacking lexical knowledge. Therefore, these students need extra cognitive support to comprehend the articles (Rose & Meyer, 2002). Extra support should include the following: presenting important concepts with pictures, audios, and videos; or teaching students to read concept maps instead of drawing.

In addition to accessible support, such as adjustable font size and color and having someone read the text aloud to them, individuals facing reading difficulties need extra cognitive support. Bottom-up and top-down are two approaches for providing cognitive support based on the reading models (Rose & Meyer, 2002). The bottom-up approach provides supports to the specific word in order to compensate readers' limitations in decoding and accessing lexical meaning; the top-down approach provides holistic maps and enhances understanding of the text. The supports of the bottom-up approach include pictures, speeches, texts, and videos for key words; adjusting attribution of text; and supplying alternative representations for the text. Earlier studies demonstrated the effect of

narration, speech, and animation on minimizing the difficulties of decoding and concentrating on the meaning of the text for the learners (Matthew, 1997; Miller, Blackstock, & Miller, 1994).

The top-down approach provides concept maps, text summarization, and background knowledge. Potelle and Rouet (2003) compared different types of content representation devices in comprehending an expository hypertext in French for 47 undergraduates. The study indicated that the hierarchical map improved comprehension for low-knowledge participants. The summarization improves comprehension by allowing readers to quickly grasp a document or preview its content (Foulds & Camacho, 2003).

Reading activity is not only a cognitive process, but also a process of motor control and visual perception. In fact, a learner should overcome three major barriers — physical, sensory and cognitive — barriers to participate in reading activities (Cook & Polgar, 2008). Thus, besides cognitive support, a learner with special needs also needs physical and sensory support. The belief description is below:

Physical Access. People with poor postural position and motor control need devices to help them maintain their body in a proper posture so that they can manipulate items appropriately. In addition, they need to be equipped with supporting devices, such as arm and wrist support systems, and have alternative access solutions, such as a mouse equipped with switches or joysticks or an infrared mouse, to navigate electronic texts and to interact with computers. While users equipped with adaptive computer devices interact with computers, they usually have to make some adjustments to the interface, such as larger icons and navigation bars, longer intervals between double clicks, and larger spaces between words and lines.

Sensory Access. The following adjustments are essential to reducing the impact of sensory difficulties for visually impaired learners: flexible interface design; enlarged font size; a font color that contrasts with the background; refreshable Braille display or a screen reader that can transfer text on demand; suitable typesetting with the adjusted space between characters, words, and paragraphs; and foreground and background colors (LoPresti, Mihailidis, & Kirsch, 2004).

Cognitive Access. Cognitive access is the core process of reading. Past experiments presented some useful strategies: marking the keywords or important ideas in the articles; reading selected text aloud; and providing an electronic dictionary, a picture symbols assistant, and a concept map (LoPresti et al., 2004). Since pictures could sometimes represent concepts more concretely than text would, they are a good assistant to reading comprehension, even for the readers with cognitive disabilities (Chen, Wu, Lin, Tasi, & Chen, 2009). Concept mapping is a kind of visualized learning strategy that could help the readers understand what they read by extracting and visualizing the key concepts from the text (Chang et al., 2002).

Though the above-mentioned strategies were proven effective by past studies, existing assistive reading software could only offer separate strategies (Chu, Li, & Chen, 2002; LoPresti et al., 2004). There has been no integrated system that can simultaneously provide physical, sensory, and cognitive access and none has especially focused on Mandarin learners with cognitive disabilities.

Four important aspects must be considered when developing a flexible, accessible, and supportive reading environment:

1. removing the three major reading barriers: physical, sensory, and cognitive.
2. providing support suited to individual needs. For example, text description of a key concept is necessary for a reader with good decoding capability but without concept knowledge; however, such support can't help a dyslexic reader.
3. keeping learning material challenging while providing support. For instance, although reading the text aloud is useful for a reader with comprehension difficulties, it will diminish the original purpose of a remedy reading program.
4. reducing professionals' workload. Although the e-learning environment allows variety in the representation of text and flexibility of the cognitive support (LoPresti et al., 2004; Rose & Meyer, 2002), certain cognitive support, such as concept maps and précis, still need to be developed by curriculum professionals. Accommodations on a case-by-case and lesson-by-lesson basis often overwhelm professionals. It would greatly ease their workload if a system in which individualized supports provided across lessons were to be developed.

Interestingly, the Universal Design for Learning (UDL), advocated by Center for Applied Special Technology (CAST), could offer thoughts that address the above-mentioned considerations. UDL consists of three principles: multi-representations, multi-engagements, and multi-expressions (Rose & Meyer, 2002). The principle of multi-representations emphasizes that learning material should be displayed in multiple representations and provide flexible cognitive supports responding to users' needs. The principle of multi-engagements stresses that learners should participate in learning activities and interact with the material their own way. The principle of multi-expressions emphasizes that learners demonstrate their learning outcomes through multiple methods. In addition, CAST believes that affectivity has great impact on learners' engagement. Past studies indicate that giving choices to learners can encourage and motivate them to participate in learning activities (Kim & Wei, 2010). Because accessible and supportive learning material is the keystone of creating UDL, it is noted that only when the material is displayed with multiple representations can learners properly interact with the learning material based on their needs and preferences. Therefore, among these three principles, developing material in multiple representations is regarded as the initial step to assisting students.

Focusing on UDL, the authors considered UDL a vital approach to creating learning material not only for special needs students but for all learners. More importantly, learning with multiple-representation material requires collaboration between material developers and instructors. Material developers must create appropriate cognitive support in advance. Then instructors can customize a reading environment for their students; moreover, learners can set up proper environments by themselves. However, no study in the past has focused on exploring the relationship among material developers, instructors, and learners, nor has a system been developed to integrate the needs of material developers, instructors, and learners.

Based on the above research and observations, this study was set to develop an individualized e-reading environment that can provide not only essential support for Mandarin readers but also a convenient interface that allows material developers to create adjustable material. In addition, an experiment was also conducted to explore the effectiveness of the system on reading comprehension for students with learning disabilities.

System development

Adopting the principle of multi-representations, this study aimed to develop an integrated, web-based e-reading system called the “TriAccess system,” which would not only provide individualized physical, sensory, and cognitive access for special needs learners but also attend to the needs of instructors and material developers.

Accessibility supports

Based on previous studies, the authors proposed 16 potential supports to address special needs learners' potential difficulties deriving from their physical, sensory, and cognitive problems. Bottom-up and top-down supports were included. Seventeen professionals in the fields of curriculum design, special education, and instruction, who were familiar with interventions for individuals with special needs, were invited to participate in a modified Delphi survey to validate these supports. They were asked to rate the importance of the 16 potential strategies. Two rounds of survey disclosed 14 items as important or very important strategies (Chen, Cko, Chen, & Chiang, 2007). As shown in Table 1, all supports except one were embedded into the system.

Table 1. Important supports for improving reading difficulties

| Accessibilities | Possible difficulties | Proposed supports |
|-----------------|---|---|
| Physical access | <ul style="list-style-type: none"> • Postural position • Motor control of upper limbs <ul style="list-style-type: none"> ▪ Gross motor ▪ Fine motor • Lacking upper limbs | <ul style="list-style-type: none"> • Adjustable interface <ul style="list-style-type: none"> ▪ Location of icons (1) ▪ Size of icons (2) ▪ Interval between icons (3) |
| Sensory access | <ul style="list-style-type: none"> • Low vision • Difficulty discriminating between foreground and background • Blind | <ul style="list-style-type: none"> • Magnifying size of text (4) • Contrasting the color of text with background (5) • Increasing spacing between words and lines (6) • Having just one sentence per line (7) |

| | | |
|------------------|---|--|
| | | <ul style="list-style-type: none"> • Reading aloud on demand (8) |
| Cognitive access | <ul style="list-style-type: none"> • Attention & perception • Word recognition • Comprehension <ul style="list-style-type: none"> ▪ Syntax ▪ Concept of keyword ▪ Background knowledge ▪ Working memory | <ul style="list-style-type: none"> • Reading aloud on demand (8) • Marking keywords or key concepts (9) • Explaining keywords in various representations (10) <ul style="list-style-type: none"> ▪ Text, voice, picture, and video • Providing background knowledge (11) • Providing graphic symbols as alternatives to text (12) * • Assisting holistic understanding <ul style="list-style-type: none"> ▪ Précis (13) ▪ Concept maps (14) |

Note: *The support was not available in the current study

TriAccess system overview

Based on the multi-user management mechanism and the multiple-representation principle, the TriAccess system is implemented through a three-tier model (client/application server/database server). Figure 1 shows the framework and the database used to store the curriculum materials and users' profiles, including specific configurations for individual learners' reading environment. The elements of an article — texts, pictures, and video clips of key concepts, summaries, and concept maps — are stored separately in the specific database. To integrate speech synthesis systems into the TriAccess system, Microsoft Agent, Visual Basic ActiveX, and Character MP3 (a popular speech-synthesized program in Taiwan) are used. The TriAccess system can "speak" the text marked by the reader. In addition, the TriAccess system adopts streaming media techniques to display video. A profile for each learner is used to manage which supports are required. The elements are assembled and displayed based on the user's profile.

Focusing on multiple-representation design, the TriAccess system consists of three user interfaces, targeting the material developer, instructor, and learner. The users' responsibilities and the functions of these interfaces are described in the following.

Material developer. Developing curriculum material requires teamwork of professionals in instruction, curriculum development, and psychology. They should prepare in advance the related cognitive supports, namely, multi-representations for explaining the key concepts, précis and background knowledge of the article, and the concept map. Then, an authorized person uploads the required cognitive supports separately using the developer's interface. The system works by organizing these supports automatically and storing them in the database. The remaining supports, including physical, sensory, and cognitive access supports, are also generated automatically. Therefore, since the above-mentioned supports are already provided by the TriAccess system, the developers need to create only the text and the required cognitive supports that the system cannot generate. The system organizes them and displays the material in response to the learners' individualized requirements, thus requiring less information-technology literacy of the instructional material developers.

Instructor. The major responsibility of authorized instructors is to set up individualized, appropriate reading environments for their students. First, students' profiles are created and then particular reading subjects are assigned. An individualized reading environment is provided by configuring proper physical and sensory access and by selecting cognitive supports based on the student's abilities, limitations, and preference. Unless the student profile is reset, the setting remains consistent in all reading material; in this way, the workload of preparing the individualized learning material can be reduced.

Learner. Learners access the specific subjects in individualized reading environment. Based on age, disabilities, cognitive capability, and computer experience, learners are placed into three categories as independent, dependent, or blind users. Users classified as independent can set up the accessibility items by themselves after logging in. Dependent users can read only in the specific environment set by the instructor. For blind users, the TriAccess system can read aloud both the article and the explanation of keywords, allowing users to access the article and the text explanation of keywords via voice by pressing hotkeys even when the computer is not equipped with a screen reader.

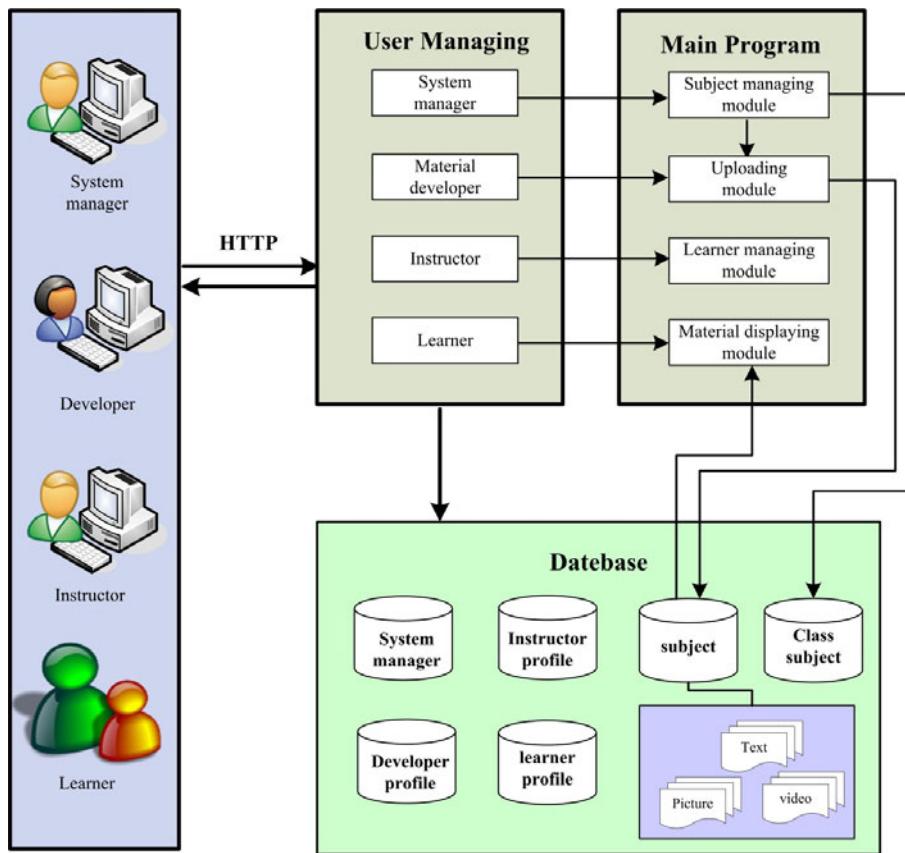


Figure 1. Framework of the TriAccess system

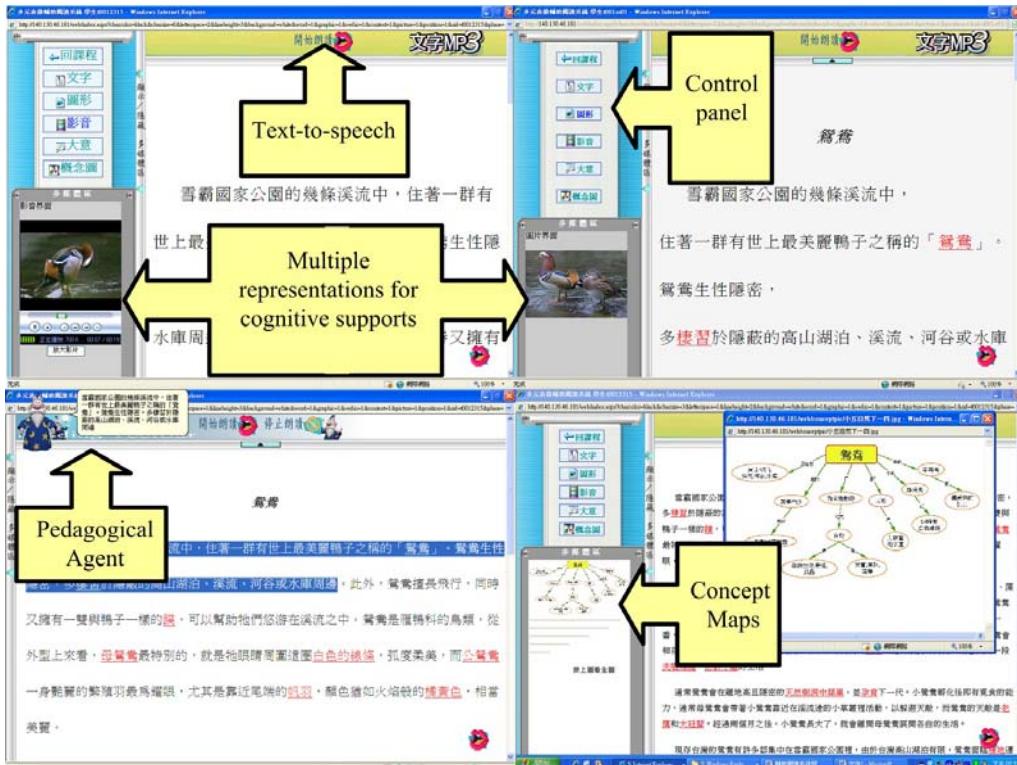


Figure 2. Four snapshots of the learner's interface, with different supports

The following demonstrates how users try to select a specific lesson after logging in to a customized environment. Some snapshots for learners without blindness are shown in Figure 2. As Figure 2 indicates, the articles are identical but displayed in different conditions and shown in four frames. The upper left frame reveals a video support of a key concept, a larger font size of text, and larger icons in the control panel. The upper right frame displays a picture support of a key concept and larger space between icons in the control panel. The lower right frame displays a concept map of the article, and the lower left frame shows that the menu bar and cognitive-support display area are disabled and that Merlin (a pedagogical agent) is activated. Friendly interface aside, pedagogical agents have high potential to better motivate learners (Dowling, 2002). The agents encourage the learner with disabilities to interact with the material actively when other visual supports are disabled.

The control panel allows learners to choose the essential support. Then they move the cursor to an underlined word. Related support would be displayed in the description area. In read-aloud assistance mode, users can mark a specific Chinese character, word, sentence, or paragraph and activate the text-to-speech program by clicking the mouth icon. Then, the computer would read out the selected text.

Usability evaluation

A preliminary usability evaluation was conducted to explore potential users' TriAccess experience (Chen et al., 2007). Thirty volunteer evaluators, consisting of five faculty members in universities, four experienced school teachers in elementary schools, thirteen special education student teachers, and eight elementary school special needs pupils, participated in the evaluation. Two versions of the five-point Likert scale regarding TriAccess system usability were developed: the special educator version and the student version. Evaluations of 22 participating educators demonstrated their positive user experience toward all the three different interfaces: material developer, instructor, and learner. Meanwhile, all eight students also expressed their satisfaction with the system.

Since the participants in the previous study (Chen et al., 2007) rated the interface after a few trials, they might not have been familiar with the system. This paper invited 40 evaluators, 30 students with learning disabilities, and 10 special education teachers to read the articles with the system, and then evaluate the usability of the interface of learner.

Revised versions of the questionnaire conducted in the previous study (Chen et al., 2007) were used. The questions in the teacher's version focused on satisfaction, remembrance, efficiency, recognizability, and learnability. Items of satisfaction rated user satisfaction with the interface design; items of remembrance examined the user could, even after they did not use the system for a period of time, still remember how to operate the system; items of efficiency tested user productivity with the system; items of recognizability tested whether the function of the system could be identified easily; and items of learnability tested whether the user could learn how to operate the system quickly. The version for students evaluated four dimensions in terms of satisfaction and learnability in addition to user preference and assistance. Items of user preference tested whether the user liked to use the interface; items of assistance examined whether users felt the system benefited them when they read the article. All the questionnaire participants rated the items on a scale of 5 to 1, with 5 meaning strongly agree and 1 meaning strongly disagree. The student version was read to student participants, who answered orally.

The results of the evaluation by both the 10 special education teachers and the 30 students demonstrated positive opinions on the usability of the reader interface. In the five-point Likert-scale questionnaire, the average score of the five dimensions for teachers — satisfaction ($M = 4.40$, $SD = .53$, $t = 2.39$, $p = .04$), remembrance ($M = 4.67$, $SD = .38$, $t = 5.48$, $p = .00$), efficiency ($M = 4.35$, $SD = .41$, $t = 2.69$, $p = .03$), recognizability ($M = 4.50$, $SD = .67$, $t = 2.37$, $p = .04$), learnability ($M = 4.60$, $SD = .32$, $t = 6.00$, $p = .00$) — were all significantly higher than 4.0. For students, the average score of the four dimensions — preference ($M = 4.56$, $SD = .77$, $t = 3.95$, $p = .00$), learnability ($M = 4.78$, $SD = .43$, $t = 9.87$, $p = .00$), satisfaction ($M = 4.63$, $SD = .55$, $t = 6.28$, $p = .00$), assistance ($M = 4.60$, $SD = .60$, $t = 5.45$, $p = .00$), were also significantly higher than 4.0. The results indicate that these participants thought that the system was user-friendly, making it easy for users to familiarize themselves with the system, could assist with comprehension, and satisfy users' need for assistance in reading.

Experiment

This experiment aimed to explore whether the cognitive support features in the TriAccess system could assist students with learning disabilities in comprehending articles. The experiment extended the authors' preliminary study (Chen, Chiang, & Ko, 2008). A two-factor within-subject experiment was conducted. The concrete research questions were as follows:

1. Is there an interaction effect between reading conditions and reading sequence?
2. Is the performance different between two reading conditions if the interaction effect is not significant?

Method

Participants. The participants were identified by the local education agent in southern Taiwan as students with learning disabilities who also had difficulty reading. Thirty students (20 boys and 10 girls) in the fifth and sixth grade participated in the experiment after acquiring their parents' consent. Their intelligence quotient (WISC-III) ranged from 70 to 109 ($M = 84.27$, $SD = 8.79$). All of them had reading difficulties and received special education service in their schools. None of them reported difficulty in vision, hearing, or motor control.

Based on the requirement of the curriculum standard in Taiwan, the students had formally begun computer and information-technology education in the third grade. Therefore, participants in the experiment had at least two years' experience using computers and the Internet. To ensure that all of them would be familiar with the TriAccess system, they were individually taught to operate the TriAccess system prior to the experiment.

Material. To exclude the effect of learning experience and demonstrate the effect of "reading for learning," the authors used the proper text developed for the authors' preliminary study (Chen et al., 2008). A panel of three experts in natural-science education assisted in developing, reviewing, and approving the experimental material — seven articles related to endemic species in Taiwan. One of them was used to familiarize the participants with TriAccess; the other six served as experimental materials. In each article, the text, keywords, or key concepts were decided first. Then, the concept map and the summary were edited. The narration, pictures, and short film were created to serve as cognitive supports for each keyword.

The six articles consisted of an average of 552 Chinese characters, ranging from 539 characters to 584 characters. The mean length of a sentence ranged from 10.9 Chinese characters to 13.8 Chinese characters, 11.8 characters on average. The structure of the six articles was consistent: all texts sequentially described creatures' living environment, characteristics, habits, and threats.

Each article was accompanied by a reading comprehension test of ten multiple-choice questions focusing on comprehension. The score ranged from 0 to 10. The higher score, the better performance. Fifty students (26 fifth grade and 24 sixth grade) without special education needs participated in evaluating the difficulty of the text by reading the article without multiple representations before they took the comprehension test. The resulting difficulty index of each comprehension test, from .65 to .85, indicated that the six tests were considered easy. Also, the Pearson's product correlation coefficient of the scores revealed a significant relationship between the scores ($p < .01$) (Chen et al., 2008).

Experiment design. A within-subjects, repeated-measures design was employed. Independent variables were reading conditions (without and with cognitive supports) and reading sequence (first, second, and third reading). Reading conditions were manipulated by randomly assigning the articles with and without cognitive supports to the participants. In order to control the impact of the related factors on reading, all articles were displayed on TriAccess system with the participants' preferred font size, color, and type. The only difference was the existence or non-existence of the cognitive supports. The articles were displayed in text mode when they were assigned without cognitive supports. On the other hand, when the articles were assigned with cognitive supports, the functions of reading aloud on demand, keywords, explaining the keywords with various representations, summary, and the concept map were enabled. Participants could use these supports during the reading process if they activated them. The authors also considered the impact of the reading sequence in this study, for the learning curve and novelty effects potentially influence the performance when children use a new system (Zhang & Zhou, 2003). Using the TriAccess system was novel for the participants. Therefore, this study aimed to explore the effect of interaction

between reading conditions and the reading sequence. However, the authors wanted to examine the main effect of the reading conditions only if the interaction effect was not significant, because the main effect of the reading sequence was not the focus of this experiment. The dependent variable was the score on the reading comprehension test.

In addition, the authors controlled the impact of experiment material on reading comprehension. As mentioned, the six articles were developed systematically to make them similar; however, they were of different content. The authors intended for each article to be read in both reading conditions to reduce the impact of experiment material. First, three pairs of the material were randomly assigned to three experiment sessions. The participants were randomly grouped into two groups. In each experiment session, first, the participants of Group One read an article without cognitive supports and those in Group Two first read the same article with cognitive supports. Then Group One read the article with multiple representations while the Group Two read the article without multiple representations. Thus, each article was read in “cognitive supports” and “text” mode randomly.

Apparatus and setting. This experiment used a laptop with a 15-inch monitor and an optical mouse to interact with the system. Chinese Character MP3 v2.0 (<http://www.iqchinese.com>), a text-to-speech software, was also installed. The participant was equipped with a headphone to enable participants to listen to the article. Due to various school locations and class schedules of the participants, the experiment was conducted individually. Students read the articles and took the tests on the TriAccess system in a quiet room in their own schools.

Experiment procedure. The authors interviewed each participant before the experiment to gather demographic information and computer experience. The second author then read the purpose of the experiment to the participant, and demonstrated how to interact with the TriAccess system. The participants practised using the system to become familiar with the reading environment. In the training sessions, the second author also helped each participant set up his/her favorite reading environment, including font size, character spacing, row spacing, and speed of reading by the text-to-speech program. These setting parameters were stored, and they served as the individual reading profile for the following experiment, for both reading conditions. Each participant needed to pass a prerequisite computer skills test in this session. Twenty nine participants passed the first time and one passed the second time.

There were three sessions in the formal experiment period. Each participant needed to read the above-mentioned six articles, three without cognitive supports and three with cognitive supports, without instruction. In a session, the participants were required to read one article with cognitive supports and the other one without cognitive supports in random sequence. When they read the text with cognitive supports, they were encouraged to use all the supports they liked and could operate the specific support as many times as they needed. Participants were also encouraged to read the article aloud when they read it without cognitive support. There was no time limit for reading the article in both reading conditions. Participants were required to take the test in which questions would be read aloud by the text-to-speech program on the TriAccess system. Their response to each question was recorded by the system. A three-minute break was given between two reading conditions.

Data analysis. SPSS 12.0 for Windows was used to analyze the data. A two-way repeated-measures ANOVA was conducted to analyze the interaction of reading conditions and reading sequence on reading comprehension. The main effect of the reading conditions was examined when the interaction of reading conditions and reading sequence on reading comprehension was not significant. The main effect of reading sequence was not examined, because the effect combined the scores from two reading conditions.

Results

The mean of test scores and standard deviations of the reading comprehension tests appear in Table 2. For these 30 participants, there was no interaction effect ($F = 0.08, p = .93$) on comprehension test scores. The main effect of reading conditions ($F = 92.18, df = 29, p = .00$) was significant. The mean of the two reading conditions demonstrated that the participants performed better when reading with cognitive supports ($M = 6.15, SD = 1.13$ for cognitive-support mode; $M = 4.30, SD = 1.20$ for non-cognitive-support mode). The practice significance was also huge ($\eta^2 = .76$). Therefore, the main effect of reading supports was confirmed.

Table 2. Descriptive statistics for two reading conditions

| Reading sequence | Reading conditions | | | |
|------------------|-------------------------|------|----------------------------|------|
| | With cognitive supports | | Without cognitive supports | |
| | Mean | SD | Mean | SD |
| First session | 6.03 | 1.88 | 4.10 | 1.76 |
| Second session | 6.60 | 1.59 | 4.87 | 1.93 |
| Third session | 5.77 | 1.61 | 3.93 | 1.76 |

Discussion and conclusions

Based on the principle of multi-representations, this paper developed an individualized and convenient e-reading system for Mandarin students with special learning needs and for curriculum professionals and teachers. The authors also examined the effectiveness of the TriAccess system on reading comprehension and explored the users' subjective perception of interacting with the TriAccess system.

In a web-based learning environment, the TriAccess system provides learners with individualized physical, sensory, and/or cognitive supports adapted to their needs. It also promises convenient and simple interfaces for material developers to prepare learning content and for the instructors to set up an individualized reading environment for their own students. The usability evaluation indicated all three interfaces were easy to use. In particular, the efficiency of the system gained the highest score. It reflected that the evaluators agreed that the TriAccess system could increase efficiency for all groups.

Effectiveness-wise, this experiment result indicated better reading performance when the students employed cognitive supports. Furthermore, this paper illustrated the effects of the cognitive supports instead of the effects of using TriAccess system on displaying digital material. In both reading conditions, the articles were displayed based on the participants' specific preferences of font size, color, and row space of the text with TriAccess system. The results demonstrated the real effectiveness of cognitive supports. However, though the study observed that the use of cognitive supports varied among participants, it didn't explore the effect of specific cognitive supports on specific individuals. Future study should further examine this aspect.

The participants' subjective attitudes also revealed a preference for the TriAccess system. Participants regarded it as a user-friendly, supportive, and interesting reading environment compared to traditional textbooks. It was observed that the TriAccess system seemed to motivate them better and, as they became more actively engaged in the learning activities, they gained better reading comprehension. However, the affective issue was not explored in this study. In conclusion, although the result of this experiment discloses greater reading comprehension when the participants read with the supportive multiple representations, some related questions have not yet been investigated. Future studies should continue exploring the issues that have not been addressed in this paper.

First, which supports will benefit the particular reader with specific disabilities needs to be explored. Teachers would like to know what specific kinds of supports might be useful for readers with certain limitations so that they can help their students to select supports accordingly. Likewise, this knowledge will help learners with disabilities identify appropriate or preferred supports and environments with less trial and error.

Second, the affective issue in reading activities should be investigated. The reasons behind their high motivation in using e-reading systems were not explored. Future study should investigate whether the reading media, selecting supports autonomously, or providing specific supports improve individuals' motivation in reading activities.

Third, the effect of integrating the teaching reading strategy with e-reading systems on reading comprehension merits further exploration. Though this paper demonstrated the effectiveness of the cognitive supports on reading comprehension, the mean score of students with disabilities on the comprehension tests was 6.15, which was not as high as the scores of the students without disabilities. The result revealed that the participants with learning disabilities still could not learn effectively by themselves through reading with cognitive supports. These learners usually lack useful reading comprehension strategies. They can't organize what they read, even though the system provides bottom-up and top-down assistance. Thus, some instruction in reading strategies may be necessary to help these students with their organization ability. For example, students may need to learn to read the concept map or to

review the key concepts of the content. Therefore, future study should integrate reading strategy into the design and explore the effect of strategy instruction, such as reading aloud after the speech is generated by text-to-speech software, or reviewing the key concepts.

Fourth, the experimental design tries to control the extraneous variables to demonstrate the effect of treatment. The design might not represent the actual situation in school, especially for exploring the reading activity. Since reading is a complex phenomenon, it is hard to control the extraneous variables and examine a single factor. That being said, the future study should adopt other research methods to explore in depth the questions for individuals with specific disabilities when reading. For example, qualitative research might be a good approach to investigate the process of improving motivation when learners read with an e-reading system, while single-subject research might be suitable for exploring the effect of integrating reading strategies and e-reading systems on reading comprehension for the individuals with disabilities.

Last but not least, the features of this e-reading system should also be enhanced in the future, at least in the following ways. First, the new version of the e-reading system should record user behavior automatically. Monitoring conditions where the supports are used by the students during the learning processes could help the learner and the teacher to determine the proper supports based on objective information. Second, it should integrate the strategies of remedy and compensation into a system to assist learners in learning reading skills and participating in reading activities through software. Many past studies demonstrated the effectiveness of software for teaching reading skills, such as concept mapping, marking key concepts, and taking notes. However, these software programs do not integrate with assisted reading features. The authors believe that an ideal supportive environment should provide both the necessary supports and essential reading strategies during reading activities.

Acknowledgements

The authors would like to thank the National Science Council of Republic of China for financially supporting this research under contract numbers 94-2614-H-415-001-F20 and 95-2614-H425-001-F20.

References

- Alliance for Technology Access (2004). *Computer and web resources for people with disabilities* (5th Ed.). Alameda, CA: Hunter House.
- Bender, W. N. (2004). *Learning disabilities: Characteristics, identification, and teaching strategies* (5th Ed.). Boston: Pearson.
- Browder, D. M., Spooner, F., Wakeman, S., Trella, K., & Baker, J. N. (2006). Aligning instruction with academic content standards: Finding the link. *Research and Practice for Persons with Severe Disabilities*, 31(4), 309–321.
- Chang, K. E., Sung, Y. T., & Chen, I. D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *Journal of Experimental Education*, 71(1), 5–23.
- Chen, M. C., Cko C. C., Chen, L. Y., & Chiang C. H. (2007). Developing and evaluating a TriAccess Reading System. *Lecture Notes in Computer Science*, 4556, 234–241.
- Chen, M. C., Chiang, C. H., & Cko C. C. (2008). The effectiveness of TriAccess Reading System on comprehending nature science text for students with learning disabilities. *Lecture Notes in Computer Science*, 5105, 747–754.
- Chen, M. C., Wu T. F., Lin, Y. L., Tasi , Y. H., & Chen, H. C.(2009). The effect of different representations on reading digital text for students with cognitive disabilities. *British Journal of Educational Technology*, 40(4), 764–770.
- Chu, C. N., Li, T. Y., & Chen, M. C. (2002). The design of an adaptive web browser for young children with reading difficulties. *Lecture Notes in Computer Science*, 2398, 189–190.
- Cook, A. M., & Polgar, J. M. (2008). *Cook and Hussey's assistive technology: Principles and practice* (3rd Ed.). Baltimore: Mosby.
- Dowling, C. (2002). Researching agent technologies in the electronic classroom: Some pedagogical issues. In P. Barker & S. A. Rebelsky (Eds.), *Proceeding of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002* (pp. 437–442). Chesapeake, VA: AACE.

- Foulds, R., & Camacho, C. (2003). Text summarization contributions to universal access. *Technology and Disability*, 15, 223–229.
- Guastello, E. F., Beasley, T. M., & Sinatra, R. C. (2000). Concept mapping effects on science content comprehension of low-achieving inner-city seventh graders. *Remedial and Special Education*, 21(6), 356–365.
- Hung, L. Y., & Huang, K. Y. (2006). Two different approaches to radical-based remedial Chinese reading for low-achieving beginning readers in primary school. *Bulletin of Special Education*, 31, 43–71.
- Kim, Y., & Wei, Q. (2010). The impact of learner attributes and learner choice in an agent-based environment. *Computer & Education*, 56(2), 505–514.
- Ledere, J. M. (2000). Reciprocal teaching of social studies in inclusive elementary classroom. *Journal of Learning Disabilities*, 33(1), 91–106.
- Lerner, J. (2006). *Learning disabilities and related disorders: Characteristics and teaching strategies* (10th Ed.). Boston: Houghton Mifflin Company.
- Lewis, R. B. (1993). *Special education technology classroom applications*. Pacific Grove, CA: Brooks Cole.
- LoPresti, E. F., Mihailidis, A., & Kirsch, N. (2004). Assistive technology for cognitive rehabilitation: State of the art. *Neuropsychological Rehabilitation*, 14, 5–39.
- Mastropieri, M. A., & Scruggs, T. E. (2002). *Effective instruction for special education* (3rd Ed.). Texas: PRO-ED.
- Matthew, K. (1997). A comparison of the influence of interactive CD-ROM storybooks and traditional print storybooks on reading comprehension. *Journal of Research on Computing in Education*, 29(3), 263–275.
- Miller, L., Blackstock, J., & Miller, R. (1994). An exploratory study into the use of CD-ROM storybooks. *Computers and Education*, 22, 187–204.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognitive and Instruction*, 1, 117–175.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327–345.
- Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Van der Bijl, C., Alant, E., & Lloyd, L. (2006). A comparison of two strategies of sight word instruction in children with mental disability. *Research in Developmental Disabilities*, 27(1), 43–55.
- Zhang, D., & Zhou, L. (2003). Enhancing e-learning with interactive multimedia. *Information Resources Management Journal*, 16(4), 1–14.

Guessing, Partial Knowledge, and Misconceptions in Multiple-Choice Tests

Paul Ngee Kiong Lau,¹ Sie Hoe Lau,¹ Kian Sam Hong,² and Hasbee Usop²

¹Universiti Teknologi MARA (UiTM), Malaysia // ²Universiti Malaysia Sarawak, Malaysia //

plnk@sarawak.uitm.edu.my // lausiehoe@sarawak.uitm.edu.my // hksam@fcs.unimas.my //

uhasbee@fcs.unimas.my

ABSTRACT

The number right (NR) method, in which students pick one option as the answer, is the conventional method for scoring multiple-choice tests that is heavily criticized for encouraging students to guess and failing to credit partial knowledge. In addition, computer technology is increasingly used in classroom assessment. This paper investigates the effect of computer-adaptive assessment software (CAAS) using the number right elimination testing (NRET) as a scoring method for multiple-choice tests. The samples were 449 Form Two students in 19 Malaysian secondary schools. These students, aged 13 to 14 years, had gone through six years of primary education and at least one year of secondary education. Drawing from the analyses performed on students' responses to the multiple-choice test in the study, we found that guessing was minimal when NRET was employed as a scoring method. NRET was able to credit partial knowledge and diagnose misconceptions. NRET scores were also consistently more reliable than the corresponding NR scores for all subtests.

Keywords

Multiple choice, NRET, Guessing, Partial knowledge, Misconceptions

Introduction

Assessment is an important component in the teaching and learning process. A large amount of classroom time is assigned to assessment-related activities. Multiple-choice (MC) tests continue to be the most common format for assessing knowledge, ability, or performance of students. Number right (NR), in which students evaluate every option and choose only one option in response to any question, is the conventional scoring method for MC. The teacher assumes that this option represents the most appropriate answer to the MC item. One mark is awarded for the correct answer, and no mark is awarded for omissions and incorrect answers. The main advantages of this format include broad content sampling, high score reliability, ease of administration and scoring, and objective scoring (Haladyna, 1994, Ben-Simon, Budescu, & Nevo, 1997). According to Olson (2005), it would cost United States 1.9 billion USD to meet testing requirement for six years if MC only were used and machine-scored. But it would cost 3.9 billion USD if both MC and open-ended items were used, and up to 5.3 billion USD if tests with written responses were hand-scored.

However, MC has been consistently criticized for having several weaknesses, such as decreased validity due to guessing and failure to credit partial knowledge (Kurz, 1999). The number of correctly answered items for a student is gathered from the number of items for which the student knows the answers and the number of items for which the student correctly guesses the answers. Hence, students could achieve higher scores due to lucky guesses. With four options per item, a student can be expected to score at least 25%. Guessing is a poor educational practice and interferes with the goal of identifying the true ability of a student from the responses to a test (Oh, 2004).

Although students may identify the wrong answer to any MC item, they can determine some of the options as definitely incorrect. This knowledge is called partial knowledge. It is believed that a student's knowledge for any MC item can be any one of full knowledge, partial knowledge, absence of knowledge, partial misconception, and full misconception, and any attempt to measure knowledge dichotomously, as in the case of the NR method, is unsatisfactory (Hutchinson, 1982). Not only is partial knowledge of students not credited, but teachers also cannot diagnose students' misunderstanding and lack of understanding in order to provide informative feedback to facilitate students' continuous learning. Despite the continuous quest for a better scoring method, none has been identified to replace the conventional NR. Holmes (2002) wrote that NR is still considered "the best of a bad lot" (p. 26).

The method of elimination of incorrect options before choosing the answer for MC items has a long history as a test-taking strategy (Shepard, 1982). However, it was never formalized as a scoring method. This study formalized it into the number right elimination testing (NRET) scoring method. NRET is a hybrid of NR and elimination testing (ET). Under NRET, students have to eliminate as many wrong options as possible and must choose one option as the

answer. However, students can choose either “correct,” or “wrong,” or “not sure” for any option. The scoring for NRET is based on the combination of NR and ET scoring methods. For any MC item with four options, one point is awarded for each wrong option eliminated correctly. However, a penalty of three points is deducted if the correct answer is eliminated. One additional point is awarded if the answer chosen is correct, and no point is given for choosing “not sure.” Both the penalty and the choice of “not sure” are included to discourage guessing. Thus, the NRET score for any MC item with four options ranges from -3 to 4. Table 1 below contains the test instruction and scoring guides for NRET.

Table 1. NRET test instruction and scoring guides

| Test Instruction |
|--|
| MUST choose ONE option as the ANSWER by using “ ✓ CORRECT .” |
| ELIMINATE option(s) that you are SURE ARE NOT THE ANSWER by using “ X WRONG .” |
| USE “ ? NOT SURE ” if you are NOT SURE of an option. |
| You have the flexibility to choose NONE (0) , ONE (1) , TWO (2) , or THREE (3) “ X WRONG ,” or “ ? NOT SURE .” |
| Scoring Guides |
| ONE (1) point awarded if the option with “ ✓ CORRECT ” is the correct answer. |
| ONE (1) point awarded for each option eliminated correctly with “ X WRONG .” |
| A PENALTY of 3 points deducted if the correct answer is eliminated with “ X WRONG .” |
| Your score will range from -3 to 4 . |

Increasingly, computer-based testing (CBT) is also being used in the classroom as computers and Internet accessibility become pervasive. Thus, studies related to CBT become important. However, most of these studies are still mainly based on NR or alternative scoring methods with complex and unfamiliar test instructions (He & Tymms, 2005).

Purpose of the study

Therefore, in this study, a computer-adaptive assessment software (CAAS) using the NRET scoring method was developed and tested in Malaysian secondary schools in 2008. The main aim was to investigate the robustness of NRET scoring method in reducing guessing and detecting partial knowledge and misconceptions. In particular, this study investigated the following four research questions:

1. Research question 1: What was the extent of guessing under NRET method?
2. Research question 2: Could NRET method detect partial knowledge and misconceptions?
3. Research question 3: What was the reliability of NRET scores as compared to NR scores?

Review of related literature

This section briefly reviews the literature related to MC tests, different scoring methods, and CBT. Major weaknesses of NR scoring are highlighted. In addition, this section also discusses various measures taken by the different scoring methods and the use of technology to improve these weaknesses.

Multiple-choice and scoring methods

Oral examination was the primary means of educational testing before the mid-nineteenth century. Written tests in the form of essay questions were introduced to replace oral examinations. Studies done in the early part of the twentieth century showed that essay tests tended to be highly subjective and unreliable in measuring students’ performance. These findings motivated educators to develop more objective educational measurements. MC was first used in 1917 for the selection and classification of military personnel for the United States Army (Ebel, 1979). Today, MC tests are the most highly regarded and widely used type of objective tests for measuring knowledge, ability, or performance (Ben-Simon et al., 1997). Traditionally, NR is the scoring method in which the number of

correct responses for a given test is taken to represent a student's ability. MC tests have been criticized for encouraging students to guess and for their inability to differentiate between various levels of knowledge (Kurz, 1999).

Students generally score higher in MC tests due to guessing. For instance, to get an average of 40% in a MC test consisting of items with four options, a student needs to know the correct answers to only 20% of the items. The student can get the other 20% through guessing correctly the answers to one quarter of the remaining 80% of the items. Hence, NR scoring fails to provide a true estimate of the knowledge of a student.

Efforts to improve the psychometric quality of MC tests focus on designing new scoring methods that reduce guessing and discriminate between different levels of knowledge. The correction for guessing method (CG) is an attempt to provide a true estimate of a student's level of knowledge by eliminating the correct responses from lucky guesses (Jaradat & Tollefson, 1988). This is based on the assumption that all incorrect responses are from guessing. CG is criticized for failing to take into account partial knowledge, and its use is rarely justified.

The recognition of partial knowledge leads to the belief that a student's level of knowledge falls on a continuum ranging from full knowledge to full misconception. Various scoring methods were proposed to credit partial knowledge: confidence weighting (CW), elimination testing (ET), subset selection testing (SST), probability measurement (PM), answer-until-correct (AUC), option weighting, item weighting, rank ordering the option, and partial ordering. All these scoring methods aim to extract information from the examinees that can provide better estimates of their abilities.

Elimination testing (ET)

ET proposed by Coombs, Milholland, and Womer (1956) is one of the most promising scoring methods that can credit partial knowledge. It requires students to pick as many incorrect options as they may identify. One mark is awarded for every incorrect option identified. However, $(k - 1)$ marks, where k is the number of options per item, are deducted if the correct option is identified as incorrect. The score for an item with four options is in the range from -3 to 3 . Thus, ET scores can help to classify a student's knowledge into full knowledge (3), partial knowledge (2 or 1), absence of knowledge (0), partial misconception, (-1 or -2) and full misconception (-3) (Bradbard & Green, 1986). ET scoring produces slightly more valid and reliable scores, but students find its test instructions to be complicated and confusing.

Traditionally, assessments have been used for comparisons among students for their performance in learning. Assessments should serve "to educate and improve student performance and not merely to audit it" (Wiggins, 1998, p. xi). NR scores for MC tests neither provide diagnostic information for teachers to identify effective classroom instruction nor can they be used by teachers to construct informative feedback for students to improve their learning. However, if the ET scoring method is used to facilitate the classification of a student's knowledge into a continuum ranging from full knowledge to full misconception, then there is a possibility to identify a student's lack of understanding and misunderstanding for any MC item. Many studies have been done to compare ET with NR. Some of these studies found that the reliability of ET scores is equal to or greater than those of NR scores and that the improvement of reliability is not statistically significant (Collet, 1971; Hakstian & Kansup, 1975; Traub & Fisher, 1977). Other results indicated that there is no loss of reliability for ET as compared to NR, and there is evidence that guessing is reduced and partial knowledge can be measured with ET (Bradbard & Green, 1986; Bradbard, Parker, & Stone, 2004; Chang, Lin, & Lin, 2007).

Technology in assessment

Advances in computer technology in the 1980s provided another opportunity for researchers to address the problem of guessing and crediting partial knowledge in MC testing. CBT has been used since the 1960s to test knowledge and problem-solving skills (Swets & Feurzeig, 1965). According to Holmes (2002), one of the earliest reported experiments with CBT was by Shuford in 1965, followed by Baker in 1968, Sibley in 1974, and Dirkzwager in 1975. From the late 1980s, computers have become more affordable and are now available in sufficient quality for use in classrooms and other educational environments. As the learning environment continues to evolve in the digital age, there is a growing interest in the development of CBT (Baucer & Anderson, 2000; Boettcher & Conrad, 1999;

Hartley & Collins-Brown, 1999; Morley, 2000). As a result, a number of innovative CBT have been proposed. Early CBT mostly adopted and modified the various existing scoring methods. Baker (1968), Dirkzwager (1975, 1993, 1996), Holmes (2002), Shuford (1965), and Sibley (1974) adopted the PM scoring method while Chambers (1990), Farrell and Lueng (2004), Klinger (1997), Paul (1994), and Rippey (1986) used a modified version of the CW scoring method. Dirkzwager (1975) started developing an interactive computer program (TestBet) based on PM. TestBet became available in 1998. Each item is presented on the screen, with a percentage slider for each alternative. Students had to slide the bar from left to right to indicate their degree of belief that an option was true or false.

However, the majority of the innovative CBT software available focus mainly on higher and further education (He & Tymms, 2005). Furthermore, CBT incorporating the CW and PM scoring methods requires a student to translate the degree of correctness of the chosen option into a fairly correct numerical scale, which requires understanding of probability. Unfortunately past studies revealed that even many adults were not able to employ probabilistic reasoning (Schwebel, 1975; Tomlinson-Keasey, 1972). Thus, this software may not be suitable for young children who are routinely tested with multiple-choice items. He and Tymms (2005) recommended developing an easy-to-use CBT, suitable for primary and secondary students.

Research methodology

The MC test is one of the most favored formats for assessing knowledge in Malaysia's education system. With the inherent weaknesses of NR scoring, educators are unable to solicit informative feedback to improve their instructional processes and facilitate students' continuous learning. If assessment is to play the role of a "powerful driver" in the teaching and learning process rather than a "terminal event," then there is a need for a more effective scoring method for MC tests.

Computer-adaptive assessment software (CAAS)

The study first developed CAAS, an online formative assessment system for MC items. Participating students could access CAAS via the website <http://caa.bestservices.com.my>. Figure 1 shows the beginning interface of CAAS. The CAAS training module consisted of five topical mathematics exercises for the first phase of training and two 40-item tests for the second phase of training. The final test used for the actual analysis contained 40 MC items. These items were adopted and modified from the mathematics items for the eighth grade students (13 years old) of the 2003 Trends in International Mathematics and Science Study (TIMSS).

Figure 1. Online formative assessment system

The use of NRET as a scoring method in this study represented a departure from normal testing routine for the students, although the method might have seemed familiar to them. The students had gone through more than seven years of formal education being assessed using the MC and NR scoring methods. There was a possibility that these students could not follow the NRET test instructions realistically and consistently. CAAS was developed in such a way that students were forced to follow NRET test instructions. If a student did not follow NRET by either not choosing one option as the answer or omitting an option, the student could not submit the solution, and a reminder would pop up telling him or her of the violation of the test instructions. Figure 2 shows such an interface of CAAS.

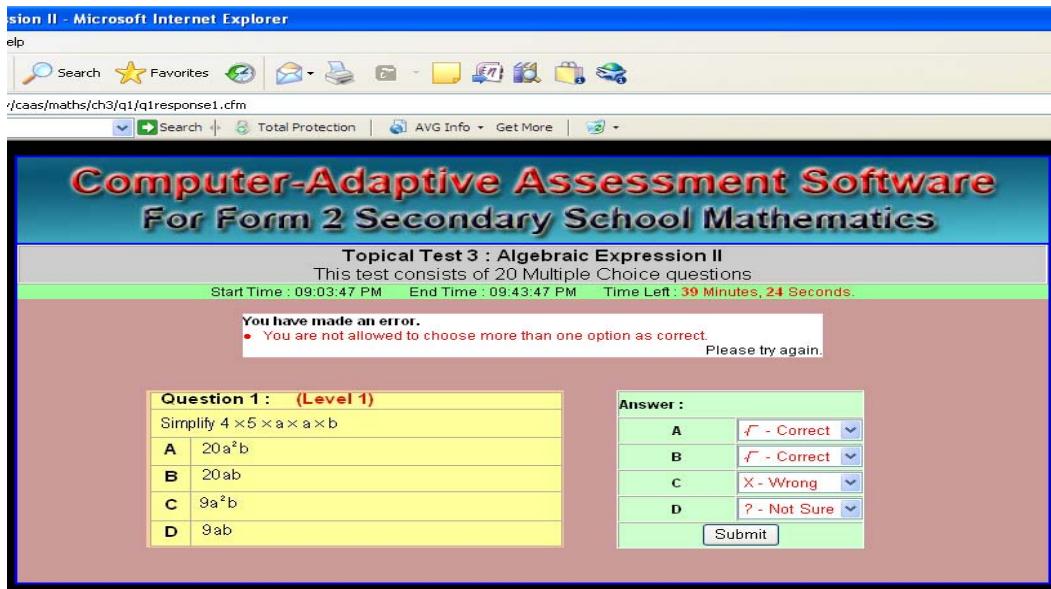


Figure 2. Not choosing one option as the answer

Feedback on student's knowledge level for each item is provided. Based on the NRET scores, students' knowledge could be classified into full knowledge (4), partial knowledge (3, 2 or 1), absence of knowledge (0), partial misconception (-1 or -2), and full misconception (-3). Figure 3 shows the partial knowledge detected by NRET, while Figure 4 shows the partial misconception detected by NRET.

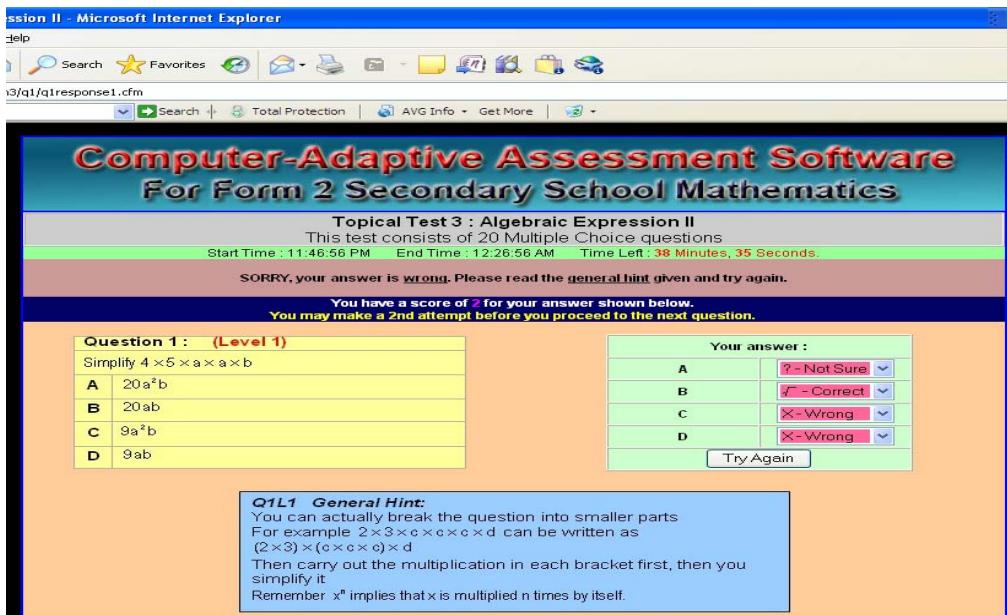


Figure 3: Partial Knowledge (a score of 2) detected by NRET

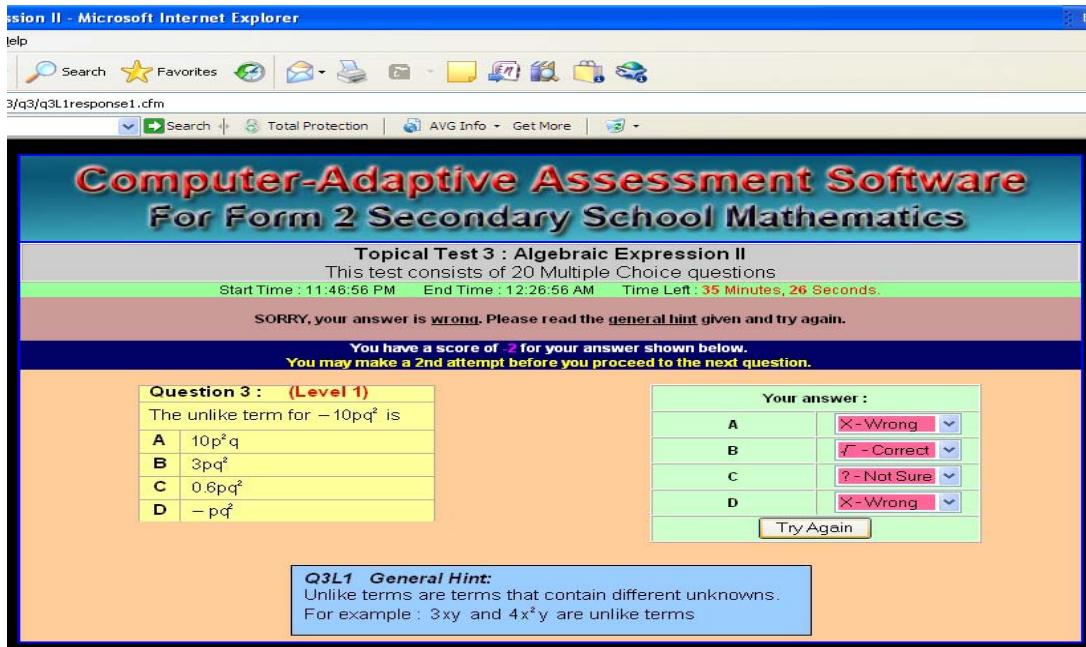


Figure 4. Partial misconception (a score of -2) detected by NRET

Subjects

Permission for the study was obtained from the Educational Planning and Research Division of the Malaysian Ministry of Education and from the Sarawak State Education Department before we met the principals of the selected secondary schools to identify students who were willing to participate. Participants were told of the purpose of the study and were not obliged to participate. If they wished to participate, their responses would remain anonymous and confidential. The students were trained to use NRET before sitting for a final MC test. A total of 449 Form Two students from 19 secondary schools in Sarawak, Malaysia, participated in this study. There were 255 female students and 194 male students, aged 13 to 14 years. They had gone through six years of primary education and at least one year of secondary education. The language of instruction for mathematics is English.

Data collection

Training was conducted before the final data was collected. The five topical exercises and the two tests were uploaded to CAAS for training, which was carried out in the school computer laboratories under the supervision of mathematic teachers appointed as research assistants for the study. However, students were permitted to complete the exercises and tests online at home if they were unable to complete them in school due to inadequate computer facilities and poor Internet connectivity.

By the end of the training sessions, the subjects were able to follow the NRET test instructions smoothly. The final test was conducted in the computer laboratory in each school using 15 laptops linked to a local server. This method of data collection was necessary to avoid interruption due to poor Internet connectivity. The subjects in each school were quarantined and sat for the final test in batches.

Data analysis

For the first research question, the extent of guessing under the NRET method was assessed based on two procedures. The first procedure, as recommended by Hambleton, Swaminathan, and Roger (1991), was to check the performance of low-ability students on the most difficult items. If guessing was minimal, the performance of the low-ability students would be closed to zero or below the chance level of 25%. The low-ability students were the

lower 30% of the sample ranked according to their NR scores (Agble, 1999). The most difficult items were three items with the lowest *p*-value (proportion of students who answer the item correctly). The focuses were on the lower ability group because they had the tendency to guess since they had only partial knowledge for most of the items (Agble, 1999). The second procedure examined the fit of the items to the two-parameter item response theory (IRT) models that assumed minimal guessing. Basically, there were three parameters involved in MC testing, namely, item difficulty, item discrimination, and guessing. The one-parameter IRT models take into consideration items having different item difficulty but assume all items to have the same item discrimination index and minimal guessing. The two-parameter IRT models take into consideration items having different item difficulty and item discrimination indices but assume minimal guessing. The three-parameter IRT models take into consideration all three parameters in MC testing. Theoretically, the three-parameter IRT models fit the data from MC tests best. However, data with minimal guessing fit the two-parameter well.

This study used BILOG-MG program for data analyses. The IRT models were the 1-Parameter Logistic (1-PL), 2-PL and 3-PL, and the response-function metrics were logistic and normal. Thus, there were a total of six possible IRT models, namely, 1-PL logistic, 1-PL normal, 2-PL logistic, 2-PL normal, 3-PL logistic, and 3-PL normal. Since the final test consisted of 40 items, the χ^2 statistics were used to assess the degree of fit of the response data to the models. If the χ^2 calculated at the 0.05 level of significance is greater than the χ^2 critical at the associated degree of freedom, then the item does not fit the model.

For the second research question, the NR and NRET scoring taxonomies were analyzed to gauge the ability of the NRET method in detecting partial knowledge and misconceptions. This procedure was employed by Chang et al. (2007) to determine the ability of ET in capturing partial knowledge. The average numbers of items for each knowledge level for the NR and NRET methods were compared. The corresponding average numbers of items would be different if the NRET method had been able to capture partial knowledge and diagnose misconceptions (Chang et al., 2007). This study further extended the procedures used by Chang et al. (2007) to include items with wrong and correct answers.

For the third research question, reliability of the NR scores and NRET scores were compared. Reliability is the degree of accuracy presented in the score and is indexed using reliability coefficient such as Cronbach's Alpha. According to Kansup (1973), there was no procedure available to test the significance difference between two alpha values within the same group. Thus, the comparison among them was made on the basis of their manifested values and the consistency of the result across tests, as used by Ma (2004) and Holmes (2002). To ensure more stable and valid values of alpha for each scoring method, simulations were carried out using the actual data by varying the subtest length. Simulation was done by randomly selecting the number of items according to the length of each subtest. A total of five simulations were conducted for each subtest length, and the average values of alpha compared. Each simulation resulted in two alpha values corresponding to the two scoring methods (NR and NRET). A higher alpha value indicates higher reliability.

Results

For the first research question, the results of the performance of the low-ability students on the most difficult items can be used to determine the extent of guessing. The three most difficult items were identified as item 16, item 26, and item 29. The responses and NR scores of the low-ability students to these three most difficult items are presented in Table 2 and Table 3.

Table 2. Responses of low-ability students to 16th, 26th, and 29th items

| Item | <i>p</i> -value | Wrong | | Correct | |
|------|-----------------|----------|------|----------|------|
| | | <i>N</i> | % | <i>N</i> | % |
| 16 | 0.388 | 121 | 88.3 | 16 | 11.7 |
| 26 | 0.370 | 110 | 80.3 | 27 | 19.7 |
| 29 | 0.430 | 120 | 87.6 | 17 | 12.4 |

As shown in Table 3, the percentages of the low-ability students' failing to answer items 16, 26, and 29 were high at 88.3%, 80.3%, and 87.6%. The assumption is that if a student guesses randomly, then the chance of getting an item

correct for a four-option MC item is 25% (one out of four). The percentages of the correct responses for all the three most difficult items were below 25% at 11.7%, 19.7%, and 12.4%. Further analysis contained in Table 3 shows that 85 out of 137 students (62.0%) had a score of 0, while 44 (32.1%) had a score of 1, and 8 (5.8%) had a score of 2. None of the low-ability students managed to get all three of the most difficult items correct. Thus, guessing was minimal with the NRET method.

Table 3. Performance of low-ability students on the three most difficult items

| NR Score for the three items | Number of students | Percentage |
|------------------------------|--------------------|------------|
| 0 | 85 | 62.0 |
| 1 | 44 | 32.1 |
| 2 | 8 | 5.8 |
| 3 | 0 | 0.0 |
| Total | 137 | 100.0 |

The number of misfit items for the one-parameter IRT models was high. Thus, the one-parameter IRT model is not appropriate, which is reflected by the large number of misfit items. However, the number of misfit items dropped significantly for the two-parameter models, which assume minimal guessing. There was no further significant decrease in the number of misfit items for the three-parameter models, which take guessing into consideration. Therefore, guessing was minimal under NRET since the data fit with the 2-parameter IRT models that assume minimal guessing. The results of the two procedures above showed that guessing was minimal under the NRET method. The results of fit of the item analyses by IRT models are listed in Table 4.

Table 4. Number of misfit items for each IRT model

| IRT Model | Number of misfit items |
|---------------|------------------------|
| 1-PL logistic | 11 |
| 1-PL normal | 13 |
| 2-PL logistic | 5 |
| 2-PL normal | 5 |
| 3-PL logistic | 3 |
| 3-PL normal | 4 |

For the second research, the results of the NR and NRET scoring taxonomies for the whole test in Table 5 show the ability of the NRET method to detect partial knowledge and misconceptions.

Table 5. NR and NRET scoring taxonomy

| Score | NR | | NRET | | | | | | FK | |
|------------|---------|-------|------|------|-------|------|------|------|------|-------|
| | Correct | Wrong | FM | PM | | NK | PK | | | |
| | | | | -3 | -2 | | 1 | 2 | | |
| Final test | 25.73 | 14.27 | 0.27 | 1.21 | 10.34 | 1.02 | 1.17 | 1.29 | 1.38 | 23.32 |

Note: FM = full misconception, PM = partial misconception, AK = absence of knowledge, PK = partial knowledge, and FK = full knowledge.

The results show that the NR scores for all items could be divided into only two categories: correct (1) and wrong (0). But the NRET scores could be divided into five different categories: FM for full misconception (-3), PM for partial misconception (-2 and -1), NK for no knowledge (0), PK for partial knowledge (1, 2, and 3) and FK for full knowledge (4), which is in line with the suggestion put forward by Ben-Simon et al. (1997).

The average number of correct items under NR was 25.73. However, the average number of items for full knowledge under NRET was only 23.32. This clearly shows that not all the correct answers under NR were based on the true knowledge of the students. Similarly, not all wrong answers under NR were due to the misconceptions or lack of knowledge on the part of students. An average of 14.27 responses were identified as wrong under NR, an average of 11.82 responses demonstrated misconceptions, and only an average of 1.02 responses were due to no knowledge. Further analyses were performed on NR and NRET scores for the correct answers and wrong answers. The results are presented in Table 6 and Table 7.

Table 6. NR and NRET scoring taxonomy for items with correct answers

| Score | NRET | | | | NR Correct |
|------------|------|------|------|-------|---------------|
| | PK | | | FK | |
| | 1 | 2 | 3 | 4 | |
| Final test | 0.54 | 0.49 | 1.38 | 23.32 | 25.73 |

Note: PK = partial knowledge, and FK = full knowledge

Table 7. NR and NRET scoring taxonomy for items with wrong answers

| Score | NRET | | | | | | NR Wrong | |
|------------|------|------|-------|------|------|------|-------------|--|
| | FM | PM | | NK | PK | | | |
| | -3 | -2 | -1 | 0 | 1 | 2 | | |
| Final test | 0.27 | 1.21 | 10.34 | 1.02 | 0.63 | 0.80 | 14.27 | |

Note: FM = full misconception, PM = partial misconception, AK = absence of knowledge, PK = partial knowledge

The results in Table 6 show that the knowledge states of the students with correct answers could be further classified into full knowledge and partial knowledge. Similarly, the results in Table 7 show that their knowledge states for wrong answers could be further categorized into full misconception, partial misconception, no knowledge, and partial knowledge under NRET. Thus, NRET could detect partial knowledge and misconceptions of the students.

For the third research question, the average reliabilities of NR scores and NRET scores were compared as in Table 8.

Table 8: Average reliability by different subtests

| Subtest length | Average reliability alpha coefficient | |
|----------------|---------------------------------------|-------|
| | NR | NRET |
| 37 | 0.907 | 0.914 |
| 35 | 0.903 | 0.911 |
| 33 | 0.899 | 0.906 |
| 31 | 0.893 | 0.900 |
| 29 | 0.887 | 0.895 |
| 27 | 0.877 | 0.887 |
| 25 | 0.868 | 0.879 |
| 23 | 0.852 | 0.864 |
| 21 | 0.847 | 0.857 |
| 20 | 0.842 | 0.854 |

The reliabilities of NRET scores were consistently slightly higher than the corresponding NR scores for all the subtests. Thus, the findings indicated that there was no loss of reliability in NRET scores as compared to NR scores, but NRET could detect partial knowledge and misconceptions.

Discussion

As more emphasis is placed on the accountability of educational institutions, the need for assessments that can provide diagnostic information for teachers to identify effective classroom instruction and help teachers construct informative feedback for students to improve learning becomes crucial. Without such focus, assessments are designed only to audit students' learning and are just "terminal events" in the teaching and learning process. MC tests are the most preferred assessments for objective measurement of students' knowledge, ability, or performance. When the dichotomous NR scoring is employed, MC tests are criticized for encouraging guessing and for their inability to discriminate between different levels of knowledge.

Over the last half a century, many scoring methods have been proposed to differentiate students' knowledge for MC tests. These methods use different test instructions and scoring mechanisms to facilitate students' responding based on their true knowledge. The reliabilities of these methods are better than those of conventional NR scoring. However, these scoring methods are not accepted as an alternative to NR due to their complex or unfamiliar test instructions and scoring rules. Among this "bad lot," ET proposed by Coombs et al. (1956), with a score ranging

from -3 to 3 for any four-option MC item, stands out for its ability to detect partial knowledge and misconceptions. They wrote as follows: "All positive scores represent some degree of partial information and all negative scores represent some degree of misinformation" (p. 35).

The study reported in this paper investigated the effect of CAAS using NRET as a scoring method for MC tests. The results were based on the analyses performed on the data collected for this study. Clearly, there are a number of limitations. First, this study used one group of students to sit for the final test under the NRET test instructions, and the NR scores were determined from these responses. This approach of calculating different scores from one test using one common test instruction had been employed by past researchers such as Kansup (1973) and Holmes (2002). It has been noted that the observed score for any item of an examinee is influenced by many variables such as guessing, testing situation, content, scoring, administration, and examinee's behavior. Through this approach, many of these variables can be held constant and the scoring errors can be minimized. However, further studies could be done with two groups of students with similar ability sitting the final test, one under NRET and the other under NR. Then, not only the reliability of NRET could be compared to NR, but the significance of a difference in reliability could also be identified. Second, we should also compare NRET to ET. Third, mathematics was the content matter and the study involved Form Two students only in Malaysia. Further studies conducted across different subjects and age groups could help to clarify the generalizability of the findings of this study. Fourth, the comparison of scoring methods should also be done using different sample distributions such as non-normal sample distributions with different values of skewness and kurtosis. Through such studies, clearer differences between the scoring methods may emerge. Last, but not least, this study was conducted using CAAS. Technology in assessment often results in unforeseen social, negative and unintended consequences. Thus, further study is needed to look into the impact on examinee's behavior regarding these issues.

Conclusion

The results showed the feasibility of adopting NRET to replace the conventional NR. First, the analyses performed by using the IRT models and on the performance of the low-ability students for the three most difficult items showed that guessing was minimal using the NRET method. These findings are consistent with findings by Swineford and Miller (1953) and Traub and Hambleton (1972), in which guessing was minimal under the penalty method. Secondly, the analyses of NR and NRET scoring taxonomies done for the whole test, on the items with correct answers and the items with wrong answers, showed that NRET could detect full knowledge, partial knowledge, absence of knowledge, partial misconception, and full misconception. These results are similar with Bradbard and Green (1986), Bradbard et al. (2004), and Chang et al. (2007), in that there is evidence that guessing is reduced and partial knowledge can be detected. Third, the analyses performed on different subtests of the final test showed that the NRET scores were consistently more reliable than the NR scores. This finding is consistent with suggestion by Ma (2004), who said that when test items were scored dichotomously, potentially useful information about individual's level of proficiency that was contained in the other response options was lost. Thus, the precision of measurement was reduced. The results of this study are similar to the results of the studies done by comparing ET with NR (Collet, 1971; Hakstian & Kansup, 1975; Traub & Fisher, 1977). Although the results of NRET and ET were comparable, NRET has an added advantage over ET. According to Jaradat and Tollefson (1988), ET instructions are confusing despite prior practice. It is conflicting where students are taught to solve for the correct answer but being assessed on their ability to identify incorrect answer. On the other hand, the test instructions of NRET resemble one of the most commonly used test strategy, in which students first eliminate the obviously incorrect options before choosing the answer.

NRET delivered through CAAS allows administrators or teachers to control the students' responses. In this case, there may have been several possible reasons for students to not employ the NRET test instructions. Firstly, they may not have understood the NRET test instructions. Secondly, they may have forgotten and lapsed into the traditional response mode. Thirdly, they may have chosen not to comply with the NRET testing mode. Regardless of the reason for noncompliance, CAAS can ensure that students comply with the required instruction. In addition, CAAS also allows speedy calculation of item scores and total scores. It also creates an opportunity to provide feedback concerning performance after each item. This can help students understand the reward or penalty associated with each response strategy. Thus, CAAS using NRET has the potential to resolve the problems of guessing and failing to detect partial knowledge and misconception that are common with the NR method.

References

- Agble, P. K. (1999). *A psychometric analysis of different scoring strategies in statistics assessment*. Unpublished doctoral dissertation, Kent State University.
- Baker, J.D. (1968). The uncertain student and the understanding computer. *Paper presented at the O.T.A.N. conference*, May 20–24, Nice.
- Baucer, J.F., & Anderson, R.S. (2000). Evaluating students' written performance in the online classroom. In R.E. Weiss, D.S. Knowlton, & B.W. Speck (Eds.), *Principles of effective teaching in the online classroom: New directions for teaching and learning* (pp. 65–71). San Francisco: Jossey-Bass.
- Ben-Simon, A., Budescu, D. V., & Nevo, N. (1997). A comparative study of measures of partial knowledge in multiple-choice tests. *Applied Psychological Measurement*, 21(1), 65–88.
- Boettcher, J.V., & Conrad, R.M. (1999). *Faculty guide for moving teaching and learning to the web*. Los Angeles, CA: League for Innovation in the Community College.
- Bradbard, D. A., & Green, S. B. (1986). Use of the Coombs elimination procedure in classroom tests. *Journal of Experimental Education*, 54, 68–72.
- Bradbard, D. A., Parker, D. F., & Stone, G. L. (2004). An alternative multiple-choice scoring procedure in a macroeconomics course. *Decision Sciences Journal of Innovative Education*, 2(1), 11–26.
- Chambers, D.B. (1990). *An evaluation of a simplified response confidence testing method for assessing partial knowledge in computer-based tests*. Unpublished doctoral dissertation, University of California, Berkeley.
- Chang, S. H., Lin, P. C. & Lin, Z. C. (2007). Measures of partial knowledge and unexpected responses in multiple-choice tests. *Educational Technology & Society*, 10(4), 95–109.
- Collet, L. S. (1971). Elimination scoring: An empirical evaluation. *Journal of Educational Measurement*, 8, 209–214.
- Coombs, C. H., Miholland, J. E., & Womer, F. B. (1956). The assessment of partial knowledge. *Educational and Psychological Measurement*, 16, 13–37.
- Dirkzwager, A. (1975). Computer-based testing with automatic scoring based on subjective probabilities. In O. Lecarme & R. Lewis (Eds.), *Computer in Education* (pp. 305–311). Amsterdam: North-Holland.
- Dirkzwager, A. (1993). A computer environment to develop valid and realistic predictions and self-assessment of knowledge with personal probabilities. In D.A. Leclercq & J.E. Bruno (Eds.), *Item banking: Interactive testing and self-assessment* (pp. 66–102). Berlin: Springer.
- Dirkzwager, A. (1996). Testing with personal probabilities: Eleven-year-olds can correctly estimate their personal probabilities. *Educational and Psychological Measurement*, 56, 957–971.
- Ebel, R. (1979). *Essentials of educational measurement* (3rd Ed.). Englewood Cliffs, NJ: Prentice Hall.
- Farrell, G., & Leung, Y.K. (2004). Innovative online assessment using confidence measurement. *Education and Information Technologies*, 9(1), 5–19.
- Hakstian, A. R. & Kansup, W. (1975). A comparison of several methods of assessing partial knowledge in multiple-choice tests: II. Testing procedures. *Journal of Educational Measurement*, 12, 231–239.
- Haladyna, T. M. (1994). *Developing and validating multiple-choice test items*. Hillsdale, NJ: Lawrence Erlbaum.
- Hambleton, R. K., Swaminathan, H., & Roger, H. J. (1991). *Fundamentals of item Response Theory*. Newburg Park, CA: Sage.
- Hartley, J.R., & Collins-Brown, E. (1999). Effective pedagogies for managing collaborative learning in on-line learning environments. *Education Technology and Society*, 2(2). Retrieved January 20, 2010, from http://www.ifets.info/journals/2_2/formal_discussion_0399.html.
- He, Q., & Tymms, P. (2005). A computer-assisted test design and diagnosis system for use by classroom teachers. *Journal of Computer Assisted Learning*, 21(6), 419–429.
- Holmes, P. (2002). *Multiple evaluations versus multiple choice as testing paradigm: Feasibility, reliability and validity in practice*. Unpublished doctoral dissertation, University of Twente, Holland.
- Hutchinson, T. P., (1982). Some theories of performance in multiple choice tests, and their implications for variants of the task. *British Journal of Mathematical and Statistical Psychology*, 35, 71–89.
- Jaradat, D., & Tollefson, N. (1988). The impact of alternative scoring procedures for multiple-choice items on test reliability, validity, and grading. *Educational and Psychological Measurement*, 48, 627–635.

- Kansup, W. (1973). *A comparison of several methods of assessing partial knowledge in multiple choice tests*. Unpublished master thesis, University of Alberta.
- Klinger, A. (1997). *Experimental validation of learning accomplishment* (Technical Report No. 970019). Pittsburgh: Frontiers in Education.
- Kurz, T. B. (1999). A review of scoring algorithms for multiple-choice tests. *Paper presented at Annual Meeting of the Southwest Educational Research Association*, January 21–23, San Antonio, TX.
- Ma, X. (2004). *An investigation of alternative approaches to scoring multiple response items on a certification examination*. Unpublished doctoral dissertation. University of Massachusetts, Amherst.
- Morley, J. (2000). Methods of assessing learning in distance education course. *Education at a Distance*, 13(1), 25–29.
- Oh, H. J. (2004, April). Reasoning test takers' guessing strategy and their understanding of formula scoring. *Paper presented at annual meeting of the American Educational Research Association (AERA)*, San Diego, CA.
- Olson, L. (2005). State test programs mushrooming as NCLB mandates as kicks in. *Education Week*, 25(13), 10–13.
- Paul, J. (1994). Alternative assessment for software engineering education. In J.L. Diaz-Herrera (Ed.), *Software Engineering Education* (pp. 463–472). New York: Springer.
- Rippey, R.M. (1986). A computer program for administering and scoring confidence tests. *Behavior Research Methods, Instruments, and Computers*, 18, 59–60.
- Schwebel, M. (1975). Formal operations in first-year college students. *The Journal of Psychology*, 91, 133–141.
- Shepard, J. F. (1982). *The Houghton Mifflin study skills handbook*. Boston, Massachusetts: Houghton Mifflin Company.
- Shuford, E.H., Jr. (1965). *Cybernetic testing*. (Report ESD-TR-65-467). Hanscom Field, Bedford, Mass: Decision Science Laboratory, L.G.
- Sibley, W.L. (1974). *An experimental implementation of computer-assisted admissible probability testing* (Research Report). Santa Monica: Rand Corporation.
- Swets, J.D., & Feurzeig, W. (1965). Computer-aided instruction. *Science*, 150(696), 572–576.
- Swineford, F., & Miller, P.M. (1953). Effects of directions regarding guessing on item statistics of a multiple-choice vocabulary test. *Journal of Educational Psychology*, 44(2), 129–139.
- Tomlinson-Keasey, C. (1972). Formal operations in females from eleven to fifty-four years of age. *Developmental Psychology*, 6, 364.
- Traub, R. E., & Fisher, C. W. (1977). On the equivalence of constructed-response and multiple-choice tests. *Applied Psychological Measurement*, 1, 355–369.
- Traub, R.E., & Hambleton, R.K. (1972). The effect of scoring instructions and degree of speedness on the validity and reliability of multiple-choice tests. *Educational and Psychological Measurement*, 32, 737–758.
- Wiggins, G. (1998). *Educative assessment: Designing assessment to inform and improve performance*. San Francisco: Jossey-Bass.

Computer-mediated Counter-Arguments and Individual Learning

Jack Shih-Chieh Hsu¹, Hsieh-Hong Huang^{2*} and Lars P. Linden³

¹Department of Information Management, National Sun Yat-sen University, Taiwan // ²Department of Information Science and Management Systems, National Taitung University, Taiwan // ³Department of Information Technology, Georgia Southern University, USA // jackhsu@mis.nsysu.edu.tw // kory@nttu.edu.tw // llinden@georgiasouthern.edu

*Corresponding author

ABSTRACT

This study explores a de-bias function for a decision support systems (DSS) that is designed to help a user avoid confirmation bias by increasing the user's learning opportunities. Grounded upon the theory of mental models, the use of DSS is viewed as involving a learning process, whereby a user is directed to build mental models so as to reduce the tendency to consider only data that supports any preexisting mental models. The results of an experiment suggest that a de-bias function, called computer-mediated counter-argument (CMCA), facilitates individual learning and improves decision satisfaction.

Keywords

Mental model, Individual learning, Computer-mediated counter-argument, De-bias, Decision support systems

Introduction

A decision support system (DSS) is a platform for learning in at least two ways. First, a DSS aids the user in extending his or her knowledge about the subject matter of the decision domain. For example, a person using a stock-selecting DSS may explore the relationships among interest rates and stock prices and, while exploring, learn knowledge about the markets. Second, a DSS may present the user with an opportunity for adjusting his or her knowledge about the decision-making process itself. In this case, the system aids a person in altering the decision-making process in an attempt to improve the outcome. An example of this second way of learning is shown by a user of a stock-selecting DSS who, having been warned of an error-producing cognitive bias known as gambler's fallacy, reacts to this warning by performing additional steps in the decision-making process, which in this case might include the reflection upon possible market outcomes, if future prices are not dependent upon historical prices. This second way of learning, an adjustment of the decision-making process, serves the decision maker by reconciling the many interlocking and incompatible beliefs about the problem domain, some of which may be based upon cognitive biases (Tversky & Kahneman, 1974). Increasingly, attention is being focused upon cognitive theories of learning and how these theories can be understood and applied in the context of DSS, a trajectory of research that is distinct from much of DSS research, which largely focuses upon the end result of decision making: decision quality (Santana, 1995). Along these lines, our focus is upon the second type of learning and how a DSS might help to adjust a user's decision-making process for the purposes of overcoming a cognitive bias known as confirmation bias.

During DSS use, cognitive biases present a perilous downside. Users, enabled by DSS, may make poor decisions and enact seemingly limited decision-making processes, often while confident that all is well. Several examples illustrate cognitive biases arising during the use of DSS. A cognitive bias called the illusion of control occurs when a user of a DSS performs a what-if analysis and displays increased confidence, yet achieves no significant performance gain (Davis & Kottemann, 1994). The cognitive bias called confirmation bias occurs when a person who is gathering information restricts attention to only data that supports a favored hypothesis. The cognitive bias known as the illusion of knowledge occurs when a user is overconfident about having access to a greater amount of information, as is easily the case with a DSS, but makes poor decisions, despite the additional information. This last cognitive bias was detected when investors who had switched from a phone-based investing method to an online trading platform became more confident, and yet recorded poor performance despite having access to more information (Barber & Odean, 2002). Attempting to counter these cognitive biases is called de-biasing (Fischhoff, 1982). In general, de-biasing techniques, such as educating users about biases, are believed to beneficially impact decision quality (Arnott, 2006) and de-biasing techniques applied to DSS have been successfully demonstrated (Bhandari, Deaves, & Hassanein, 2008). To investigate the embedding of a de-biasing technique into the design of a DSS we adopt mental-model theory.

The theory of mental models explains how people perform certain activities, for example, information processing. When a decision maker confronts a large amount of information, a mental model aids in the filtering of that information. The mental model fulfills a gatekeeper role for the mind by preventing unrelated and unimportant information from being consciously considered. However, along with the beneficial effects that are attributed to mental models, there are also error-producing effects. Inadequate mental models may block important information. The ability of a decision maker to change mental models is viewable as a learning process, one that may be enabled by DSS. That a person can change existing mental models is important, given that mental models may have a negative impact during decision making (Vandenbosch & Higgins, 1994). Through the mental-model theory, we seek to explain how an embedded de-bias function can induce a change in a person's mental model, conceptualized as a type of learning, and, as a result, overcome confirmation bias. Mental models provide a theoretical explanation upon which to ground a DSS function for reducing confirmation bias.

The consideration of mental models is not new to DSS research and can be recognized as a type of cognitive skill. The use of DSS is attributed to the development of cognitive skills. Intelligent tutoring tools, for example, provide the user with an opportunity to obtain cognitive skills, such as the ability to learn via self-explanation (Patel, Kinshuk, & Russell, 2000). Technology-enriched learning includes meta-learning activities, such as monitoring and sequencing one's own learning (Sinitsa, 2000). The cognitive skill being considered in this research is the learning effect of mental-model reforming. We seek to understand if the design of a de-biasing function embedded in a DSS increases the cognitive skill of the type of learning associated with mental-model reforming.

The de-bias function being explored aims to eliminate cognitive bias through the introduction of counter-arguments. The de-bias function that is embedded in a DSS is designed to present counter-arguments to the user so as to eliminate bias through a learning process. We seek to understand if users, having read counter-arguments, report to have built mental models and express a lower confidence level, as compared with users without counter-arguments. Being aware of the need to lessen the impact of confounding variables from the environment, we conducted an experiment to investigate this embedded de-bias function.

In the following sections, we discuss the concepts of mental-model theory, confirmation bias, and learning. We propose a research model grounded upon mental-model theory and state our hypotheses. We discuss the research design and the procedures of the experiment. We analyze the data and discuss the results in light of their implication to DSS design and individual learning.

Theoretical background and hypotheses

Mental models and confirmation bias

While explaining the nature of thought, Craik (1943) describes the modeling capabilities of the human mind, whereby "small-scale models" of external entities are manipulated by the mind to consider past events and to anticipate future events. These mental models are psychological representations of the environment and consist of cognitive constructs and their relationships (Holyoak, 1984). As such, mental models serve a range of purposes: when designing, mental models help a person to describe the composition of entities; when observing a system, mental models help a person to explain the current states of unseen variables; and when considering the future, mental models help a person to predict events and suitable responses to these events (Rouse & Morris, 1986). Mental models, as structured patterns, aid in classifying information and aid in simulating a set of outcomes based upon varying conditions (Cannon-Bowers, Salas, & Converse, 1993). An activity such as information processing suggests that mental models are relevant to DSS design.

One example illustrates the association of mental models with cognitive biases. In the stock market, a decision maker needs to predict a future price in order to make the purchasing decision. The decision maker possesses a mental model of the impacts of various factors upon price. This mental model helps to guide the search for information required for predicting price and abandon information that is not important to the decision. If abundant amounts of information are available, the mental model serves as a filter that helps the decision maker forego consideration of inconsequential information. A mental-model-based information search can reduce processing time and minimize information overload. These benefits, however, which coexist with a mental model, may lead the decision maker to

neglect, or perceive as unimportant, information critical to the analysis. Mental models contribute to a cognitive bias called confirmation bias.

Wason (1960) defined the phenomenon of confirmation bias as people's tendency to seek evidence in support of their assumptions instead of searching for evidence that challenges their assumptions. When a person contemplates concepts or verifies assumptions, a variety of biased tendencies have been described. A person may tend to treat evidence supporting existing beliefs more favorably than is objectively appropriate. A person confronted with a body of information may "see" exactly the effect that the person set out to find (Russo, Medvec, & Meloy, 1996). A person may tend to pass along information that is congruent to their beliefs (Heath, 1996). In short, the problem is that "decision-makers seek confirmatory evidence and do not search for disconfirming information" (Arnott, 2006, p. 60). The consequence of overweighting some evidence and underweighting other evidence is reaching a wrong conclusion. Confirmation bias is observed to be an exceptionally problematic aspect of human reasoning (Nickerson, 1998). How then might the design of a DSS compensate for this cognitive bias?

Mental models and learning

Based on mental models and learning theory (e.g., Gagne, 1977; Norman 1982; Piaget 1954), Vandenbosch and Higgins (1994) argue that mental models are closely connected to learning. Mental models direct the information-gathering process and limit the process. Reflexively, the information gathered has the potential to change mental models, confirm them, enhance them, change them, and reinforce them. Vandenbosch and Higgins (1994) state that the gathered information impacts mental models through two different processes: mental-model maintenance and mental-model building. Mental-model maintenance occurs when new information fits easily into existing mental models. Mental-model building occurs when new information from newly perceived situations and environments provokes a person to change existing mental models.

The connection between mental models and learning is amplified by Vandebosch and Higgins (1994) with their comparison of mental models to theories found in the management discipline. First among these theories is Argyris and Schön's (1978) organizational learning theory of single-loop and double-loop learning. Single-loop learning represents the learning that takes place when existing routines guide the problem-solving process. Double-loop learning represents the challenging of existing routines and the need to restructure existing norms and assumptions. These two ways of learning are similar to the learning that is explained by the use of mental models. In terms of mental models, single-loop learning maps to mental-model maintenance, both of which pertain to the confirmation and disconfirmation of new information when compared to an existing, stable theory. Double-loop learning maps to mental-model building, both of which represent a potential change in foundations of how learning is achieved.

Vandenbosch and Higgins (1994) also draw parallels between mental models and another theory, March's (1991) theory of exploitation and exploration learning. Exploitation is the act of enhancing performance by improving current practices. Exploration is the act of enhancing performance by introducing new practices. Exploitation represents mental-model maintenance as one attempts to reinforce current mental models by adding new constructs or building new links between existing constructs. Exploration represents the mental-model building process as one attempts to reconstruct the model completely (March, 1991). These management theories, because they include both the learning of individuals and the collective of individuals (i.e., the organization as a whole) may not be fully applicable to the individual learning considered in this research. However, in as much as they are found to model individual learning, these theories are useful because they are congruent with our understanding of the maintenance and building of mental models.

Of the two approaches described, mental-model maintenance and mental-model building, the one more likely to happen is mental-model maintenance (e.g., Quinn, 1980; Kiesler & Sproull, 1982; Grønhaug & Falkenberg, 1989). Acceptance of mental-model maintenance as a form of learning is high because it can be routinely confirmed through a person's past experience. Vandenbosch and Higgins (1994) write how, in contrast, mental-model building implies that decision makers face uncertainty. The theory of mental-model building has been tested through empirical evidence. People tend to make decisions or take in part or all action based on their mental models. The mental-model-building form of learning requires stimulus such as feedback from actions and information that contrasts with current thinking. DSS could be designed to induce mental-model building by injecting a particular type of information. Therefore, we attempt to identify the presence of mental-model building by introducing the stimulus of

computer-mediated counter-arguments (CMCA). A counter-argument is information contrasting with what the decision maker previously hypothesized, and computer-mediated counter-argument is defined as opposite information provided by a DSS during the decision-making process. In addition, in our study, we also attempt to determine whether individuals using a system with counter-argument report a higher level of satisfaction. The research model (Figure 1) shows the relationship between these factors.

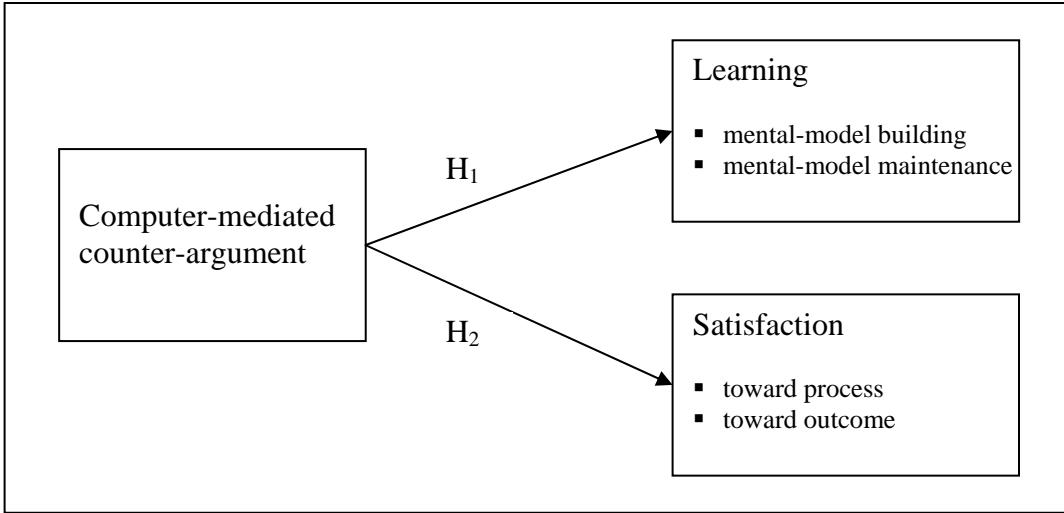


Figure 1. Research model

From a cognitive perspective, existing mental models determine the ideal construction and concept framing process. Due to cognitive-based biases, people tend to read information that fits into their current mental model. The information-selection process allows them to gather evidence that can support their existing mindset. Although the received information either strengthens or weakens the current mental model (Vandenbosch & Higgins, 1996), theory indicates that once one's belief is formed, it is not easy to change if there is no additional stimulus (Hoch & Deighton, 1989). Since subjects in the group without computer-mediated counter-argument tend to maintain their current mental model rather than building a new mental model, therefore, we hypothesize the following:

H1a: The score of mental-model maintenance for subjects who use DSS without computer-mediated counter-argument is higher than for those who use DSS with computer-mediated counter-argument.

On the other hand, people question their existing mindset when they receive contradictory evidence (Vandenbosch & Higgins, 1996). Counter-arguments provided by the system give contradictory evidence that challenges decision makers' perspectives and provides a chance for them to reevaluate their current mental model. Under this setting, people are more likely to abandon their current mental model or take appropriate steps to improve their mental model. Therefore, subjects in the group with computer-mediated counter-argument tend to reform their mental model, rather than maintain their mental model. We hypothesize the following:

H1b: The score of mental-model building for subjects who use DSS with computer-mediated counter-argument is higher than for those who use DSS without computer-mediated counter-argument.

Decision satisfaction can be separated into process satisfaction and outcome satisfaction (Green & Taber, 1980). Process satisfaction refers to users' perceived efficiency and fluency of using a decision support system to facilitate decision making. On the other hand, outcome satisfaction refers to users' perceived effectiveness and expected performance toward the decision that has been made (Sanders & Courtney, 1985). For subjects in the group with CMCA, the existence of counter-argument challenges their prior assumptions, and this process may lead to negative emotional outcomes. Decision makers have to spend more time on making a decision and the fluency of the decision-making process is low. Negative moods and increased fatigue are caused by interruptions (Schonpflug & Battman, 1988; Zohar, 1999). Therefore, decision makers tend to rate satisfaction with the decision-making outcome as low, even if the interruption is helpful (Speier, Vessey, & Valacich, 2003). However, different results are expected toward

the outcome satisfaction. Although the counter-arguments interrupt the decision-making process by allowing decision makers to consider both positive and negative opinions suggested by the system, the decision maker is predicted to report satisfaction toward their decision. Therefore, we hypothesize the following:

H2a: Subjects who use DSS with computer-mediated counter-argument will feel less satisfied about the decision-making process.

H2b: Subjects who use DSS with computer-mediated counter-argument will feel more satisfied about the decision-making outcome.

Research methods

An experiment was conducted to test the listed hypotheses. Participants were assigned randomly into two groups, with/without computer-mediated counter-argument. They were asked to accomplish a stock-investment task using the DSS provided by the authors. A stock-investment task was selected because DSS are intensively adopted to support stock investment decision making to cope with high uncertainty and process data in large quantities.

In general stock-investment decision-making tasks, subjects are allowed to invest certain amounts of money in one or more selected investing targets. Real decision making is more complex, and the final investment portfolio may contain several combinations of investments. Since this study focuses on understanding learning effects only, a simplified investment environment was constructed. That is, people were requested to invest their money in two companies belonging to different industries instead of among all available companies in the open market.

The stock-investment task used in this paper was modified from Melone, McGuire, Hinson, and Yee (1993). Subjects were requested to put the given amount of money into two different stocks. To maximize their return, subjects allocated that money into two different options (a company in the financial holding industry or high-tech industry) based on whatever subset they deemed appropriate. DSS with “what if” analysis was provided as a decision aid. The system provided basic market information and current market status (with news format). Basic market information included the name of company, capital, business scope, expected ROI in the next five years, and stock price for the past five years. Current market status included positive and negative information about these two companies and the industries in which they belong. Information was summarized from articles published in investing-related journals and magazines.

Although the investment information presented in this experiment was collected from real financial news, events, and reports, the content was reviewed and modified by two business-school PhD candidates. We removed the names and brand names that appeared in the original news report to avoid any inappropriate connection between the experiment and real-world events. Further, we refined the wording of news titles and news content. The length of the presented news headlines and story bodies were condensed to 10–20 words and 150–200 words, respectively.

The content validity was then assured by eight financial professionals, including professors, experienced investors, senior managers in the high-tech industry, and accountants. News with ambiguous titles or content was dropped or modified. The direction (positive or negative) and importance of each piece of news and the accompanying headline was determined based on opinions provided by the above eight financial professionals.

Measurement

User satisfaction includes two major dimensions: satisfaction toward the system and toward the decision making (Sanders and Courtney, 1985). Since this study focused on the new function, namely counter-argument provided by the DSS, several questions were used to understand users’ attitude toward the DSS. A total of three questions adopted from Sanders and Courtney (1985) were used to measure subjects’ satisfaction with the DSS. On the other hand, three questions adopted from Green and Taber (1980) were used to measure satisfaction toward the decision-making process.

Learning in this study refers to, with regard to DSS use, participants' using the information to challenge their assumptions or using the information to maintain their initial beliefs. A total of six questions obtained from (Vandenbosch & Higgins, 1996) were used to measure learning effects: three for mental-model maintenance and three for mental-model building. The measurement uses a seven-point Likert scale ranging from 1 (strong disagree) to 7 (strong agree).

Manipulation check

Since we wanted to manipulate the effect of with and without counter-argument, we asked subjects whether or not they were aware of the pop-up window showing counter-argument information during the experiment. We found that 90 subjects in the group with CMCA were aware and only two subjects were not; and that 86 subjects in the group without CMCA were not aware (or could not recall) and only nine subjects were. The result of subjects' awareness is shown in Table 1.

Table 1. Results of subjects' awareness

| | | Awareness | | | Total number |
|---------------------------------------|-----|-----------|----|------------------|--------------|
| | | Yes | No | Could not recall | |
| Computer-mediated Counter-argument | w/o | 9 | 71 | 15 | 95 |
| | w | 90 | 1 | 1 | 92 |
| Total number | | 99 | 72 | 16 | 187 |

According to a reading-behavior study by using eye-tracking machines, in a two-column condition, people tend to focus on content that appears on the left-hand side first. Since the information for these two companies appeared in a two-column format, we randomized the location of the messages to avoid the possible confounding effect. As shown in Table 2, the location of the news and whether or not there was CMCA are independent.

Table 2. Results of subjects in different information displayed location

| Location of displayed information | Computer-mediated counter-argument | | Total number of subjects |
|--|------------------------------------|----|--------------------------|
| | w/o | w | |
| IT-related news on the left-hand side | 51 | 45 | 96 |
| IT-related news on the right-hand side | 44 | 47 | 91 |
| Total number of subjects | 95 | 92 | 187 |

Table 3. DSS functions (Hung et al., 2007)

| Menu | Function |
|---------------------|---|
| Market summary | Detailed information on the stock market, including stock quotes, change, day's range, volume, etc. |
| Industry quotes | Detailed information on the specific industries. |
| Company information | Provided information on six companies used for trading in the experiment, including basic information, historical prices, news, and streaming charts. |
| Technical analysis | Analyzed technical aspects, such as the moving average, relative strength index, and stochastic line indicators. |
| Recent news | Provided three categories of news: political, international, and financial. |
| Calculator | Calculated investment returns and more. |

System

A GUI-based DSS, with 4GL programming language and a relational database management system, was developed for the experiment. The system was able to support Simon's decision processes, which allow people to perform intelligence, design, and choice activities. Actual design activities followed Hung, Ku, Liang, and Lee's recommendations (2007), and the detail system tools are shown in Table 3. Furthermore, the purpose of this study was to understand how users' information-reading behaviors affected their final decision. To support this, a total of 5 to 15 positive and 5 to 15 negative message titles about each company were displayed on the main page for the non-information-loading group. After users clicked on the message title, the system popped up an "always on top" screen,

which contained detailed message content. To record the exact reading time for each message, this pop-up “always on top” window could only be closed by clicking the “close” function. Finally, for experiment group, after they clicked the “close” button, another window containing a reversed direction message popped up.



Figure 2. Screenshots of the experimental system

Procedure

Before the experiment started, each participant read and signed a consent form describing the purpose of the research. Subjects were told that any decision they made was independent from their academic performance and that

the experiment process was anonymous. All experiments were held by the same facilitator to reduce possibilities of bias.

In the first stage, the experimenter gave a five-minute introduction of two investment options (a financial holding company and a high-tech company). After the brief introduction, subjects were asked to make a quick decision first. We consider the decision at this stage as prior belief. Next, what-if analysis tools and the current status of each market were shown on the screen. Positive and negative information was represented by a brief title only. Participants had to click on the title to obtain detailed information. They could use what-if analysis tools to calculate possible returns according to their understanding of the market and the provided information. In this stage, the subjects' information acquisition behaviors, including reading sequence and time, were recorded by the system automatically. When participants felt comfortable making their final decision, they pressed a button on the screen, and the system led them to the final decision screen. After participants made their final decision, they were asked to provide their level of confidence toward their final decision and answer predefined questions. Finally, they were thanked for their participation.

Pilot tests

The experiment was pilot-tested twice with master and senior students from the MIS department. After the first pilot test, with eight volunteer students, we refined the process flow and divided the company information page into two pages: one for company information and the other for the preference test. In addition, we invited three students studying human-computer interfaces to review our experiment system and provide us with some functional and interface-design comments. Based on these comments, we improved our system and made its operation smoother and more user-friendly. The second pilot test was conducted with fifteen students. We found no revisions were needed after the second pilot test, so this version was considered final.

Demographics

A total of 187 subjects were recruited from part-time MBA programs of seven different universities and one academic institution. Since our predefined task was stock investment, we focused on subjects with stock or other investing experience. Five of the six subjects had experience with stock investment, while one of the six had no experience with stock investment but had other investment experience. The average age was 33.1 years old with a standard deviation of 5.6. The eldest subject was 52 years old, and the youngest was 22 years old.

For those subjects in the with-CMCA group, one out of five had never invested in the stock market, and the average age was 33.19 years old with a standard deviation of 4.843. The eldest subject was 48 years old, and the youngest was 22 years old. For those subjects in the without-CMCA group, one out of five had never invested in stock, and the average age was 33.01 years old with a standard deviation of 6.325. The eldest subject was 52 years old, and the youngest was 22 years old.

Table 4. Validity and reliability

| Variables | Items | Factor loading | | | |
|---|-----------------------------|----------------|----------|----------|---------|
| | | Factor 1 | Factor 2 | Factor 3 | Factor4 |
| Mental-model maintenance (Cronbach's alpha = 0.74) | Support my viewpoint | 0.86 | | | |
| | Support action taking | 0.77 | | | |
| | Reinforce belief | 0.70 | | | |
| Mental-model building (Cronbach's alpha = 0.70) | Challenge my viewpoint | | 0.82 | | |
| | Criticize my cognition | | 0.81 | | |
| | Reinvestigate my assumption | | 0.73 | | |
| Satisfaction process (Cronbach's alpha = 0.87) | Process efficiency | | | 0.79 | |
| | Process satisfaction | | | 0.69 | |
| | Process clarity | | | 0.68 | |
| Satisfaction decision (Cronbach's alpha = 0.95) | Happy with the decision | | | | 0.83 |
| | Pleased with the decision | | | | 0.82 |
| | Satisfied with the decision | | | | 0.82 |

Data analysis

Item reliability, convergent validity, and discriminant validity tests are often used to test the robustness of measurement. Individual item reliability can be examined by observing the factor loading of each item. Convergent validity should be assured when multiple indicators are used to measure one construct. Convergent validity can be examined by bivariate correlation analyses and reliability of questions. Discriminant validity focuses on testing whether the measures of constructs are different from one another. It can be assessed by testing whether the correlation between pairs of construct are below the threshold value of 0.90. As shown in Table 4, given that all criteria were met, validity and reliability of our measurement are assured in our study.

Hypotheses testing

There are four hypotheses in this category, based on different decision outcomes: mental-model maintenance, mental-model building, decision-process satisfaction, and outcome satisfaction. In this study, MANOVA was selected because the correlation matrix showed a moderate- to high-level correlation among variables and to control an experiment-wide error rate resulting from conducting multiple ANOVA (Hair, Black, Babin, Anderson & Tatham, 2006). In Table 5, descriptive analysis and the correlation among variables were also provided for the validation of assumption. “Mean” represents the averaged response toward each construct. “Skewness” and “kurtosis” represent the extent to which our sample fit into the normality assumption. The MANOVA test result in Table 6 shows that, except for process satisfaction, CMCA has effects on each dependent variable when confirmation bias is controlled.

Table 5. Descriptive analysis and correlation matrix

| | Descriptive analysis | | | | Correlation matrix | | | |
|-----------------------------|----------------------|----------|----------|----------|--------------------|-------------|-------------|-------------|
| | Mean | Std dev. | Skewness | Kurtosis | MMM | MMB | SATP | SATD |
| Mental-model maintenance | 4.87 | 0.78 | -0.03 | 1.00 | 1.00 | | | |
| Mental-model building | 5.16 | 0.87 | -0.09 | -0.16 | -0.09 | 1.00 | | |
| Satisfaction toward process | 4.97 | 0.93 | -0.71 | 1.51 | 0.31 | 0.45 | 1.00 | |
| Satisfaction toward outcome | 5.02 | 0.97 | -0.33 | 0.51 | 0.32 | 0.46 | 0.78 | 1.00 |

Table 6. Effect of CMCA on learning and satisfaction

| Source | Dependent variable | F | Sig. |
|-----------|-----------------------------|--------------|-------------|
| Intercept | MM building | 6109.007 | .000 |
| | MM maintenance | 7370.924 | .000 |
| | Process satisfaction | 5246.856 | .000 |
| | Outcome satisfaction | 4992.449 | .000 |
| CMCA | MM building | 4.711 | .031 |
| | MM maintenance | 7.912 | .005 |
| | Process satisfaction | 2.388 | .124 |
| | Outcome satisfaction | 5.868 | .016 |

The hypothesis that the score of mental-model maintenance for subjects who use DSS without CMCA will be higher than for those who use DSS with CMCA is supported. As Table 7 indicates, for subjects in the group without CMCA, the average-rated mental-model maintenance was 5.03 (out of 7). For subjects in the group with CMCA, the average-rated mental-model maintenance score was 4.71 (out of 7). The significant result indicates that CMCA can effectively reduce mental-model maintenance.

The hypothesis that the score of mental-model building for subjects who use DSS with CMCA will be higher than for those who use DSS without CMCA is supported. In addition, as Table 7 shows, for subjects in the group without CMCA, the average-rated mental-model building was 4.96 (out of 7); for subjects in the group with CMCA, the average-rated mental-model building was 5.36 (out of 7).

The MONOVA result shows that counter-argument has an effect on outcome satisfaction but not on process satisfaction. The hypothesis that subjects who use DSS with CMCA will feel more satisfied toward the system is not

supported. However, significant difference can be found between these two groups and H2, and thus, the satisfaction toward decision-making outcome is supported. The average scores and standard deviation for process and outcome satisfaction for each group are provided in Table 7.

Table 7. Analytical results of mental-model maintenance

| | Computer-mediated counter-argument | |
|--------------------------|------------------------------------|-------------|
| | w/o | w |
| Mental-model maintenance | 5.03 (0.88) | 4.71 (0.64) |
| Mental-model building | 4.96 (0.81) | 5.36 (0.87) |
| Process satisfaction | 4.88 (0.90) | 5.07 (0.96) |
| Outcome satisfaction | 4.86 (0.89) | 5.19 (1.02) |

Discussion

One of our hypotheses (H2) was not supported. Contrary to our expectation that subjects provided with CMCA would be interrupted by the popping-up of counter-argument during the decision-making process and would tend to feel less satisfied, subjects in the group with counter-argument felt even more satisfied than the group without counter-argument. A couple of reasons may have caused this to happen. First, decision-process satisfaction is measured by a subject's perceived efficiency, satisfaction, and clearness toward the decision-making process. In general, decision makers do not read all the information provided by the system. Accompanied by the experimenter's introduction, the decision-making process is very easy to understand and follow. The measurement may not exactly reflect the actual problem with counter-argument. Therefore, there is no significant difference between these two groups. There is a need to develop another instrument to understand subject's satisfaction toward the decision-making process.

Second, we measured the subjects' perceived process satisfaction and perceived outcome satisfaction together, after the decision had been made. Although the questionnaire focused on the decision-making process, the decision makers' perceived satisfaction toward the process may have been affected by their perceived satisfaction toward the decision outcome. Therefore, future research should reconsider the most appropriate time for collecting process satisfaction.

Conclusion

Because human beings have bounded rationality, our decision-making process is largely affected by cognitive biases (Tversky & Kahneman, 1974). This study focused on the introduction of a de-bias function, which may reduce the negative consequences of confirmation bias in the computer-supported decision-making context. The relationships between counter-argument and decision outcomes were hypothesized. After collecting data from 187 experienced financial investors, three out of four proposed hypotheses were supported (the relationship between the providing of CMCA and process satisfaction was not supported). We successfully showed how the de-bias function can trigger individual learning by leading decision makers to challenge their current mental models. This result contributes to academics and practitioners in several ways.

Implications for research

First, our results indicate that decision support systems with specific functions can stimulate learning by reforming decision makers' mental model. How users can learn through the use of information systems has long been ignored by information system research (Alavi & Leidner, 2001). In this study, we focused on two different approaches to learning: mental-model maintenance and mental-model building. However, biases emerged while decision makers attempt to maintain current mental models. This research shows that with pre-design functions, users may challenge their initial beliefs and reform their mental models.

Second, there was no difference with regard to process satisfaction with or without counter-argument. This indicates that although a new function may be useful to minimize overconfidence and to guide the decision maker to balance

their decision, people may be against the de-bias function in the decision-making process. The existence of counter-argument may interrupt decision makers' information reading and downgrade their processing fluency. This implies that, although counter-argument allows decision makers to challenge their assumptions, an unwanted emotional side effect may also appear. Future research may explore other emotional consequences of using de-biasing functions embedded in computer-supported decision-making tools.

Implications for practice

Learning takes place when learners enhance their current mental models or build new mental models. It is vital to build an environment to help individuals to avoid cognitive bias as well as to support the building and reforming of mental models. Computer-supported learning is popular and broadly adopted by the contemporary education system. In addition to providing various types of knowledge in a more efficient or effective manner through the support of information technology, approaches such as counter thinking, which provide stimulus for learners to challenge their existing mental model, should be adopted by educators or embedded into the learning system.

For a decision-support-systems designer, de-bias functions should be included in the system design. Our study shows that systems should provide certain functions to prevent bias or support learning. However, designers should also note that some unwanted effects may be caused by those de-bias functions. For example, in our study, we provided one piece of counter-argument when subjects read a piece of news. The pop-up counter-argument may lead to negative emotions, for example, annoyance or confusion. Too many counter-arguments or too much information may increase the effect of information overload. Therefore, the frequency and timing used to provide counter-argument is important. System designers should also consider the type of task and other factors.

Limitations and future research

This study is not without limitations. First, a between-group comparison was conducted and subjective evaluation of perceived mental-model building was used to evaluate the learning effect. Some may argue that although this approach may capture part of the learning effect, a pre- and post-experiment comparison should be made in order to truly reflect how counter-argument changes the mental model. Therefore, further research should extend this study by employing within-individual experiment and examining the change of one's mental model objectively. Second, this was a cross-sectional study and the existence of counter-argument was unexpected by the subjects. People may get used to the appearance of counter-argument and ignore it automatically (e.g., Handzic & Tolhurst, 2002). Future research may extend this study by conducting a longitudinal study to examine the above assumption. Third, only a financial decision-making task was used. Since confirmation bias can be applied to various decision-making tasks and counter-argument may serve as a useful tool for eliminating it, future research should extend the proposed concept to other decision-making areas to test the effect of counter-argument on decision-making. Fourth, only one de-biasing function was included in this study. One bias may be eliminated through different approaches under different settings. Future research should include other de-biasing functions and compare the effect of different de-biasing functions under different settings. Lastly, although we have illustrated the effect of confirmation bias and methods to prevent it, we did so with a simple experiment design. Future studies may examine this issue in more complicated circumstances, such as a group decision-making or social decision-making process.

References

- Alavi, M., & Leidner, D. E. (2001). Research commentary: Technology-mediated learning: A call for greater depth and breadth of research. *Information System Research, 12*(1), 1–10.
- Argyris, C., & Schön, D. (1978). *Organizational Learning: A Theory of Action Approach*. Boston, MA: Addison-Wesley.
- Arnott, D. (2006). Cognitive biases and decision support systems development: A design science approach. *Information Systems Journal, 16*(1), 55–78.
- Barber, B. M., & Odean, T. (2002). Online investors: Do the slow die first? *The Review of Financial Studies, 15*(S2), 455–487.
- Bhandari, G., Deaves, R., & Hassanein, K. (2008). Debiasing investors with decision support systems: An experimental investigation. *Decision Support Systems, 46*(1), 399–410.

- Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. N. J. Castellan (ed.), *Individual and group decision making*. Hillsdale, NJ: Lawrence Erlbaum Associates, 221–246.
- Craik, K. J. W. (1943). *The Nature of Explanation*. Cambridge, UK: Cambridge University Press.
- Davis, F. D., & Kottemann, J. E. (1994). User perceptions of decision support effectiveness: Two production planning experiments. *Decision Sciences*, 25(1), 57–76.
- Festinger, L. A. (1957). *Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Fischhoff, B. (1982) Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, Cambridge University Press, 422–444.
- Gagne, R. M. 1977. *The condition of learning*, New York: Holt, Rinehart and Winston.
- Green, S. G., & Taber, T. D. (1980). The effects of three social decision schemes on decision group process. *Organizational Behavior and Human Performance*, 25(1), 97–106.
- Grønhaug, K., & Falkenberg, J. S. (1989). Exploring strategy perceptions in changing environments. *Journal of Management Studies*, 26(4), 349–359.
- Hair, J. F., Black, B., Anderson, R. E., Tatham, R. L., & Black, W. C. (2006). *Multivariate data analysis*, 6th ed. Englewood Cliffs, NJ: Prentice Hall.
- Handzic, M., & Tolhurst, D. (2002). Evaluating an interactive learning environment in management education. *Educational Technology & Society*, 5(3), 113–122.
- Heath, C. (1996). Do People Prefer to Pass Along Good or Bad News? Valence and Relevance of News as Predictors of Transmission Propensity. *Organizational Behavior and Human Decision Processes*, 68(2), 79–94.
- Hoch, S. J., & Deighton, J. (1989). Managing What Consumers Learn from Experience. *Journal of Marketing*, 53(2), 1–20.
- Holyoak, K. J. (1984). Analogical thinking and human intelligence. In *Advances in the psychology of human intelligence*, R. J. Sternberg (Eds.), Hillsdale, NJ: Erlbaum, 199–223.
- Hung, S. Y., Ku, Y. C., Liang, T. P., & Lee, C. J. (2007). Regret avoidance as a measure of DSS success: An exploratory study. *Decision Support Systems*, 42(4), 2093–2106.
- Kiesler, S. & Sproull, L. (1982). Managerial response to changing environments: Perspectives on problem sensing from social cognition. *Administrative Science Quarterly*, 27(4), 548–570.
- March, P. (1991). How to people with a mild/moderate mental handicap conceptualise physical illness and its cause. *British Journal of Mental Subnormality*, 37(73), 80–91.
- Melone, N. P., McGuire, T. W., Hinson, G. B., & Yee, K. Y. (1993). The effect of decision support systems on managerial performance and decision confidence. *Proceedings of the 26th Hawaii International Conference on System Sciences*, 482–489.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Norman, D. A. (1982). *Learning and memory*. San Francisco: W. H. Freeman and Company.
- Patel, A., Kinshuk, & Russell, D. (2000). Intelligent tutoring tools for cognitive skill acquisition in life long learning. *Educational Technology & Society*, 3(1), 32–40.
- Piaget, J. (1954). *The construction of reality in the child*, New York: Basic Books.
- Quinn, J. B. (1980). *Strategies for change: Logical incrementalism*, Homewood, IL: Richard D Irwin.
- Rouse, W. & Morris, N. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological bulletin*, 100(3), 349–363.
- Russo, J. E., Medvec, V. H., & Meloy, M. G. (1996). The distortion of information during decisions. *Organizational Behavior and Human Decision Processes*, 66(1), 102–110.
- Sanders, G. L., & Courtney, J. F. (1985). A field study of organizational factors influencing DSS success. *MIS Quarterly*, 9(1), 77–93.
- Santana, M. (1995). Managerial learning: a neglected dimension in decision support systems. *Proceedings of the 28th Annual Hawaii International Conference on System Sciences*, 82–91.
- Schonpflug, W., & Battmann, W. (1988). The costs and benefits of coping. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition, and health*, New York: Wiley, 699–713.

- Sinitsa, K. (2000). Learning individually: A life-long perspective. *Educational Technology & Society*, 3(1), 17–23.
- Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity and information presentation on computer-supported decision-making performance. *Decision Sciences*, 34(4), 771–797.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Vandenbosch, B., & Higgins, C. (1994). The measurement of learning from executive information systems. *Proceedings of the Twenty-Seventh Hawaii International Conference on System Science*, Hawaii, 572–585.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, 12(3), 129–140.
- Zohar, D. (1999). When things go wrong: The effect of daily work hassles on effort, exertion, and negative mood. *Journal of Occupational and Organizational Psychology*, 72(3), 265–283.

Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems

Yi-Hsuan Lee^{1*}, Yi-Chuan Hsieh² and Chia-Ning Hsu¹

¹Department of Business Administration, National Central University, Zhongli, Taiwan // ²Department of Applied Foreign Languages, Ching Yun University, Zhongli, Taiwan // yihsuanlee@mgt.ncu.edu.tw // ychsieh@cyu.edu.tw // ning0910@hotmail.com

*corresponding author

ABSTRACT

This study intends to investigate factors affecting business employees' behavioral intentions to use the e-learning system. Combining the innovation diffusion theory (IDT) with the technology acceptance model (TAM), the present study proposes an extended technology acceptance model. The proposed model was tested with data collected from 552 business employees using the e-learning system in Taiwan. The results show that five perceptions of innovation characteristics significantly influenced employees' e-learning system behavioral intention. The effects of the compatibility, complexity, relative advantage, and trialability on the perceived usefulness are significant. In addition, the effective of the complexity, relative advantage, trialability, and complexity on the perceived ease of use have a significant influence. Empirical results also provide strong support for the integrative approach. The findings suggest an extended model of TAM for the acceptance of the e-learning system, which can help organization decision makers in planning, evaluating and executing the use of e-learning systems.

Keywords

E-learning system, Technology Acceptance Model (TAM), Innovation Diffusion, Etheory (IDT), Employee training, Structural equation modeling, System adoption, End-users' perception

Introduction

To maintain competitiveness and keep a highly-trained and educated workforce, organizations have invested considerable amount of time and resources in e-learning as a supplement to traditional types of training, because it can be simultaneously implemented company –wide, achieve immediacy, consistency and convenience, and is associated with higher profits and lower turnover, thus playing a significant role in training and development (DeRouin, Fritzche & Salas, 2005).

Many studies have discussed the benefits of e-learning applications (Ong, Lai, & Wang, 2004; Piccoli, Ahmad, & Ives, 2001). But, despite increased usage, underutilization remains a problem (Moore & Benbasat, 1991; Johansen & Swigart, 1996; Ong et al., 2004). Therefore, if learners fail to use learning systems, the benefits of such systems will not be achievable (Pituch & Lee, 2006; McFarland & Hamilton, 2006). Researchers and practitioners alike strive to find answers to the problem by investigating individuals' decisions on whether or not to adopt e-learning systems that appear to promise substantial benefits (McFarland & Hamilton, 2006; Xu & Yuan, 2009; Venkatesh, Morris, Davis, & Davis, 2003). To this end, studies of user perceptions and of understanding factors involved in promoting effective use of these systems (Mun & Hwang, 2003) have become increasingly essential to improve understanding and prediction of acceptance and utilization (Lau & Woods, 2008). Prior empirical studies strived to explicate the determinants and mechanisms of users' adoption decisions on the basis of the technology acceptance model (TAM) (Davis, Bagozzi, & Warshaw, 1989; Taylor & Todd, 1995; Venkatesh & Davis, 2000) with the conviction that the adoption process influences successful use of particular technology systems (Karahanna, Straub, & Chervany, 1999; Liao, Palvia, & Chen, 2009).

This study contributes to the TAM literature by examining the relationships between the innovation diffusion theory and TAM variables in the same model. We propose to examine the effects of motivational determinants on TAM constructs using IDT as a background theory. Thus, we employed five factors: relative advantage, compatibility, complexity, trialability and observability as determinants of perceived usefulness (PU), perceived ease of use (PEU) and behavioral intention to use (BI). This empirical study could be useful for developing and testing theories related to e-learning system acceptance, as well as to practitioners for understanding strategies for designing and promoting e-learning systems.

E-learning and TAM

The TAM has been widely used as the theoretical basis for many empirical studies of user technology acceptance and has partially contributed to understanding users' acceptance of information systems (IS)/information technology (IT) (Taylor & Todd, 1995; Venkatesh & Davis, 2000). Our research shows that many studies focus on the acceptance by students in educational institutions (Chang & Tung, 2008; Pituch & Lee, 2006), but acceptance within organizations is rarely covered, and very few studies have adopted the TAM as a model for explaining the use of an e-learning system designed and provided by organizations. TAM could be useful in predicting end-users' acceptance of an e-learning system in organizations (Davis et al., 1989; Arbaugh, 2002; Wu, Tsai, Chen, & Wu, 2006); additionally, existing antecedents of the technology acceptance intention in the TAM model do not sufficiently reflect the e-learning system end users' acceptance within organizations (Ong et al., 2004; Lau & Woods, 2008).

In our model, employees' PU of the e-learning systems is defined as the perception of degrees of improvement in learning because of adoption of such a system. PEU of the e-learning systems is the users' perception of the ease of adopting e-learning systems. We made assumptions that the more end-users who perceive usefulness of the e-learning systems within an organization, the more positive their acceptance of e-learning systems, consequently increasing their chances for future usage of the e-learning systems (Arbaugh & Duray, 2002; Pituch & Lee, 2006). Furthermore, technology acceptance is determined by behavioral intention to use (Ajzen & Fishbein, 1980). Therefore, within an organizational context adoption of an e-learning system is a positive function of the intention (BI) to accept the systems.

Theoretical background

Although much research supports the TAM as an excellent model to explain the acceptance of IS/IT, it is questionable whether the model can be applied to analyze every instance of IS/IT adoption and implementation. Many empirical studies recommend integrating TAM with other theories (e.g. IDT, or DeLone & McLean's IS success model) to cope with rapid changes in IS/IT, and improve specificity and explanatory power (Carter & Be'langer, 2005; Legris, Ingham, & Colerette, 2003).

TAM and IDT are similar in some constructs and complement each another to examine the adoption of IS/IT. Researchers indicate that the constructs employed in TAM are fundamentally a subset of perceived innovation characteristics; thus, the integration of these two theories could provide an even stronger model than either standing alone (Wu & Wang, 2005; Chen, Gillenson, & Sherrell, 2002). Past studies integrated the two theories, providing good results (Sigala, Airey, Jones, & Lockwood, 2000; Chen et al, 2002).

This study employs two major theoretical paradigms—the TAM (Gefen, 2004; Talyor & Todd, 1995; Davis et al., 1989) and IDT (Roger, 1995; Moore & Benbasat, 1991). After reviewing literature on technology acceptance, we synthesized the major theories and empirical research, then proposed a model that blended key constructs involved in e-learning system acceptance and intention to use the e-learning systems.

Five constructs of innovative characteristics, PEU, and usefulness and intention to use the e-learning system, were taken from the TAM and IDT. With appropriate modifications, our proposed model could successfully be generalized to acceptance within an organizational context.

The Technology Acceptance Model (TAM)

The TAM was derived to apply to any specific domain of human-computer interactions (Davis et al., 1989). The TAM asserts that two salient beliefs —PU and PEU—determine technology acceptance and are the key antecedents of behavioral intentions to use information technology. The first belief, PU was the degree to which an individual believes that a particular system would enhance job performance within an organizational context (Davis et al., 1989). PEU, the second key belief, was the degree to which an individual believes that using a particular system would be free of effort (Davis et al., 1989). In addition, the model indicated that system usage was indirectly affected by both PEU and PU.

Many researchers have conducted empirical studies to examine the explanatory power of the TAM, which produced relatively consistent results on the acceptance behavior of IT end users (Igbaria, Zinatelli, Cragg, & Cavaye, 1997; Venkatesh & Davis, 2000; Horton, Buck, Waterson, & Clegg, 2001). Researchers have agreed that TAM is valid in predicting the individual acceptance of numerous systems (Chin & Todd, 1995; Segars & Grover, 1993). In summary, TAM provided an explanation of the determinants of technology acceptance that enables explanation of user behavior across a wide scope of end-user information technologies and user populations (Davis et al, 1989).

Innovation Diffusion Theory (IDT)

Research on the diffusion of innovation has been widely applied in disciplines such as education, sociology, communication, agriculture, marketing, and information technology, etc (Rogers, 1995; Karahanna, et al., 1999; Agarwal, Sambamurthy, & Stair, 2000). An innovation is “an idea, practice, or object that is perceived as new by an individual or another unit of adoption” (Rogers, 1995, p. 11). Diffusion, on the other hand, is “the process by which an innovation is communicated through certain channels over time among the members of a social system” (Rogers, 1995, p. 5). Therefore, the IDT theory argues that “potential users make decisions to adopt or reject an innovation based on beliefs that they form about the innovation” (Agarwal, 2000, p. 90).

IDT includes five significant innovation characteristics: relative advantage, compatibility, complexity, and triability and observability. Relative advantage is defined as the degree to which an innovation is considered as being better than the idea it replaced. This construct is found to be one of the best predictors of the adoption of an innovation. Compatibility refers to the degree to which innovation is regarded as being consistent with the potential end-users' existing values, prior experiences, and needs. Complexity is the end-users' perceived level of difficulty in understanding innovations and their ease of use. Trialability refers to the degree to which innovations can be tested on a limited basis. Observability is the degree to which the results of innovations can be visible by other people. These characteristics are used to explain end-user adoption of innovations and the decision-making process.

Theoretically, the diffusion of an innovation perspective does not have any explicit relation with the TAM, but both share some key constructs. It was found that the relative advantage construct in IDT is similar to the notion of the PU in TAM, and the complexity construct in IDT captures the PEU in the technology acceptance model, although the sign is the opposite (Moore & Benbasat, 1991). Additionally, in terms of the complexity construct, TAM and IDT propose that the formation of users' intention is partially determined by how difficult the innovation is to understand or use (Davis, et al., 1989; Rogers, 1995). In other words, the less complex something is to use, the more likely an individual is to accept it. Compatibility is associated with the fit of a technology with prior experiences, while the ability to try and observe are associated with the availability of opportunities for relevant experiences. These constructs relate to prior technology experience or opportunities for experiencing the technology under consideration. Compatibility, and the ability to try and observe can be treated as external variables, which directly affect the constructs in the technology acceptance model. After the initial adoption, the effects of these three constructs could be diminished with continuous experience and reduced over time (Karahanna et al., 1999).

Thus far, numerous studies successfully integrated IDT into TAM to investigate users' technology acceptance behavior (Hardgrave, Davis, & Riemenschneider, 2003; Wu & Wang, 2005; Chang & Tung, 2008). Few have attempted to examine all IDT characteristics with the integration of TAM. In this research, we improve TAM by combining IDT characteristics, adding compatibility, complexity, relative advantage, and the ability to try and observe as additional research constructs to increase the credibility and effectiveness of the study.

Research model and hypotheses

We propose an integrated theoretical framework, which blends TAM and IDT theories. The research model holds that the five innovative characteristics (compatibility, complexity, relative advantage, ability to try and observe) exert an important effect on the employees' PU, PEU and intention to use e-learning systems. We thus tested the validity and applicability of the proposed model based on the following hypotheses.

Compatibility

Agarwal and Prasad (1999) asserted a positive relationship between an individual's prior compatible experiences and the new information technology acceptance. They found that the extent of prior experience with similar technologies was positively associated with an ease of use belief about an information technology innovation. Moreover, Chau and Hu (2001) reported that the effect of compatibility was found to be significant only in relation to PU. Later, Wu and Wang (2005) and Chang and Tung (2008a) confirmed that compatibility had a significant positive and direct effect on PU and the behavioral intention. Likewise, prior studies have investigated compatibility from different aspects, resulting in support for its impact on PU, PEU and intention to use (Hardgrave et al., 2003). Based upon the preceding research, the following hypotheses were proposed:

H1-1: Compatibility had a positive effect on PU of the e-learning system.

H1-2: Compatibility had a positive effect on PEU of the e-learning system.

H1-3: Compatibility had a positive effect on behavioral intention to use the e-learning system.

Complexity

Empirical studies provided evidence indicating that complexity had a significantly negative effect on the intention to use (Shih, 2007; Lee, 2007). Additionally, a negative relationship between complexity and PU was also revealed in a study conducted by Hardgrave, et al. (2003). Similarly, empirical research has also shown that the more complex the end users perceived the e-learning system as being, the lower the users' intention to use the system (Lin, 2006). Thus, based on the aforementioned studies, we proposed the following hypotheses:

H2-1: Complexity negatively affected PU of the e-learning system.

H2-2: Complexity negatively affected PEU of the e-learning system.

H2-3: Complexity negatively affected behavioral intention to use the e-learning system.

Relative advantages

Research consistently found that the perceived relative advantages positively affected the users' intention to use the system across different participants (Shih, 2007; Lee, 2007). However, in TAM and IDT research, the relationships among relative advantages, PU, and PEU had seldom been studied with the only one study revealed that when the users perceived higher relative advantages, they perceived a higher level of usefulness of the systems. Accordingly, we hypothesized:

H3-1: The relative advantages had a positive effect on PU of the e-learning system.

H3-2: The relative advantages had a positive effect on PEU of the e-learning system.

H3-3: The relative advantages had a positive effect on behavioral intentions to use the e-learning system.

Observability

Using different methodologies and involving different participants from many fields, some studies found that observability had a positive impact on the users' attitude toward the system and intention to use the system (Lee, 2007). Also in line with previous studies combining TAM and IDT, when the employees perceived the systems as being easier to be observed or described, they tended to perceive the systems more useful and easier to use (Huang 2004; Yang, 2007). Therefore, we proposed that observability would have a positive effect on PU, PEU, and behavioral intention to use the e-learning system. The following hypotheses tested these assumptions:

H4-1: Observability had a positive effect on PU of the e-learning system.

H4-2: Observability had a positive effect on PEU of the e-learning system.

H4-3: Observability had a positive effect on behavioral intention to use the e-learning system.

Trialability

Some studies have empirically tested in understanding the association between trialability and the intention to use the system (Lee, 2007). They found that trialability had a positive effect on the intention to use the system. However, limited research has been conducted to investigate the relationship among trialability, PU, PEU, and behavioral intentions to use the systems. There was only one research reported that when the users perceived higher trialability, they perceived higher levels of usefulness, and ease of use of the system (Yang, 2007). Accordingly, we tested the following hypotheses:

H5-1: Trialability had a positive effect on PU of the e-learning system.

H5-2: Trialability had a positive effect on PEU of the e-learning system.

H5-3: Trialability had a positive effect on behavioral intention to use the e-learning system.

PEU

PEU is the degree to which an individual believes that using a particular system would be free of effort (Davis et al., 1989). Information system researchers have indicated that PEU has a positive effect on the end-users' behavioral intention and PU to use the systems (Chin & Todd, 1995). Thus, we hypothesized:

H6-1: PEU had a positive effect on the PU of the e-learning system.

PU

PU is the degree to which an individual believes that a particular system would enhance his or her job performance within an organizational context (Davis et al., 1989). Information system researchers have investigated TAM, and asserted that PU was valid in predicting the individual's acceptance of various systems (Venkatesh & Davis, 2000). Previous studies discovered that PU positively affected the users' behavioral intention to use systems (Chin & Todd, 1995). Therefore, we hypothesized:

H6-2: PU will have a positive effect on the behavioral intention to use the e-learning system.

Table 1: Demographics of the respondents

| Demographics | Number | % |
|---------------------------|--------|------|
| Gender | | |
| Female | 260 | 47.1 |
| Male | 292 | 52.9 |
| Age | | |
| <29 | 320 | 58.0 |
| 30-39 | 155 | 28.1 |
| 40-49 | 49 | 8.9 |
| > 50 | 28 | 5.1 |
| Education | | |
| High school | 13 | 2.4 |
| College/University degree | 308 | 55.8 |
| Master degree | 224 | 40.6 |
| Doctoral degree | 7 | 1.3 |
| Experience with computers | | |
| < 1 year | 120 | 21.7 |
| 1 to 3 years | 173 | 31.3 |
| 3 to 6 years | 83 | 15.0 |
| 6 to 9 years | 54 | 9.8 |
| > 9 years | 122 | 22.1 |

Research methodology

The subjects and the procedure

This study utilized a web-based and mailed survey to collect data for quantitative testing of the research model. Because of the lack of a reliable sampling frame, it proved difficult to conduct a random sampling for all the end-users in the organizations using e-learning systems in Taiwan. Thus, in this study we adopted a non-random sampling technique (i.e. convenience sampling) to collect the sample data. To generalize results, we gathered sample data from the five largest e-learning systems using industries (Chan, 2005), including manufacturing, finance, marketing and service, information technology, and government agencies in Taiwan, and randomly selected 15 firms that provide an e-learning training system for employees (three in each industry). Of the 736 mailed and electronic questionnaires, 566 were completed and returned. Sample demographic information is depicted in Table 1.

Measures

To ensure content validity of the scales, the items chosen for the constructs were adapted from previous research to ensure content validity. The questionnaire consisted of three parts. The first part was based on nominal scales and the rest are 5-point Likert scales. Part 1 of the questionnaire was based on IDT including compatibility (CPA), complexity (CPL), relative advantages (ADV), observability (OB), and trialability (TRI). The above items were adapted from the previous studies (Davis et al., 1989; Moore & Benbasat, 1991; Taylor & Todd, 1995; Karahanna et al., 1999), containing 18 items.

Part 2 of the questionnaire was based on the constructs of PU, PEU, BI in the TAM model and was adapted from the measurement defined by Davis et al. (1989) and Venkatesh & Davis (2000), containing 12 items for the above constructs. Part 3 of the questionnaire was to collect the interviewees' basic demographic data, such as gender, educational level, work experience, prior experience using computers, etc.

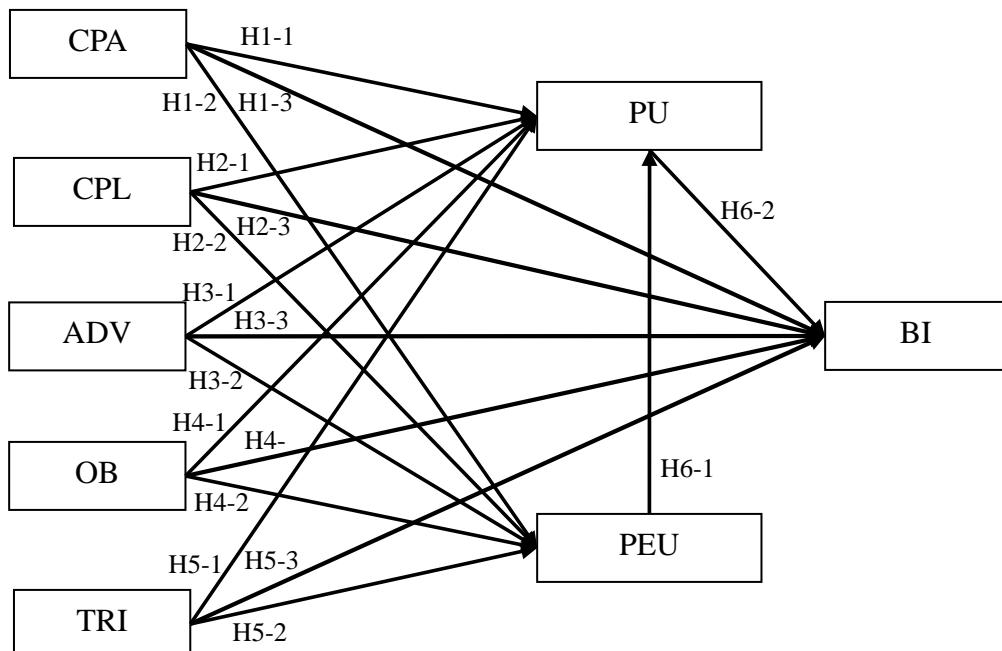


Figure 1. Proposed research model.

Results

Instrument validation

Two confirmatory factor analyses (CFA) were computed using AMOS 6.0 to test the measurement models. The model-fit measures were used to assess the model's overall goodness of fit (χ^2 / df , GFI, NFI, CFI, RMSEA) and values all exceeded their respective common acceptance levels (Hair, Black, Babin, Anderson, & Tatham, 2006). This showed that the measurement model exhibited a fairly good fit with the collected data (Table 2).

Table 2: Fit indices for endogenous and exogenous measurement models

| Goodness-of-fit measure | Recommended value | Endogenous measurement model | Exogenous measurement model |
|-------------------------|-------------------|------------------------------|-----------------------------|
| χ^2 / df | ≤ 3.00 | 1.764 | 1.977 |
| GFI | ≥ 0.90 | 0.979 | 0.958 |
| AGFI | ≥ 0.90 | 0.960 | 0.936 |
| NFI | ≥ 0.90 | 0.967 | 0.967 |
| CFI | ≥ 0.90 | 0.983 | 0.983 |
| RMSEA | ≤ 0.05 | 0.037 | 0.042 |

Table 3: Convergent validity

| Constructs/Factors | Indicators | Standardized loadings (>0.707) | Reliability (R^2) (>0.50) | Composite reliability (>0.70) | Average variance extracted (>0.50) |
|--------------------|------------|------------------------------------|-----------------------------------|-----------------------------------|--|
| CPA | CPA1 | .807 | .652 | .849 | .585 |
| | CPA2 | .720 | .518 | | |
| | CPA3 | .791 | .626 | | |
| | CPA4 | .739 | .547 | | |
| CPL | CPL1 | .854 | .730 | .906 | .764 |
| | CPL2 | .918 | .842 | | |
| | CPL3 | .848 | .719 | | |
| ADV | ADV1 | .777 | .604 | .926 | .716 |
| | ADV2 | .812 | .660 | | |
| | ADV3 | .876 | .768 | | |
| | ADV4 | .905 | .819 | | |
| | ADV5 | .854 | .729 | | |
| OB | OB1 | .744 | .554 | .857 | .670 |
| | OB2 | .953 | .908 | | |
| | OB3 | .740 | .547 | | |
| TRI | TRI1 | .790 | .624 | .827 | .615 |
| | TRI2 | .838 | .703 | | |
| | TRI3 | .720 | .518 | | |
| PU | PU1 | .847 | .717 | .854 | .663 |
| | PU2 | .870 | .757 | | |
| | PU3 | .717 | .514 | | |
| PEU | PEU2 | .769 | .591 | .841 | .570 |
| | PEU3 | .766 | .587 | | |
| | PEU4 | .703 | .494 | | |
| BI | BI1 | .675 | .456 | .915 | .686 |
| | BI2 | .773 | .598 | | |
| | BI3 | .935 | .874 | | |
| | BI4 | .879 | .773 | | |
| | BI5 | .854 | .729 | | |

Convergent validity of scale items was estimated by reliability, composite reliability, and average variance extracted (Fornell & Larcker, 1981). The standardized CFA loadings for all scale items exceeded the minimum loading criterion of 0.70, and the composite reliabilities of all factors also exceeded the recommended 0.70 level. In addition,

the average variance-extracted values were all above the threshold value of 0.50 (Hair, et al., 2006). Hence all three conditions for convergent validity were met for the four measurement models (See Table 3).

Discriminant validity was obtained by comparing the shared variance between factors with the average variance extracted from the individual factors (Fornell & Larcker, 1981). This analysis showed that the shared variances between factors were less than the average variance extracted for the individual factors. Hence, discriminant validity was assured (see Table 4). To sum up, the four measurement models reached satisfactory levels of reliability, convergent validity and discriminant validity.

Table 4: Discriminant validity

| Construct | Interconstruct correlations | | | | | | | |
|-----------|-----------------------------|-------|-------|-------|-------|-------|-------|-------|
| | BI | PU | PEU | CPA | CPL | ADV | OB | TRI |
| BI | 0.828 | | | | | | | |
| PU | 0.353 | 0.814 | | | | | | |
| PEU | 0.286 | 0.229 | 0.755 | | | | | |
| CPA | 0.466 | 0.401 | 0.253 | 0.765 | | | | |
| CPL | 0.180 | 0.068 | 0.572 | 0.210 | 0.874 | | | |
| ADV | 0.368 | 0.375 | 0.269 | 0.624 | 0.138 | 0.846 | | |
| OB | 0.138 | 0.052 | 0.094 | 0.123 | 0.061 | 0.138 | 0.819 | |
| TRI | 0.228 | 0.095 | 0.271 | 0.240 | 0.240 | 0.185 | 0.203 | 0.784 |

Note. Diagonals represent the square root of average variance extracted, and the other matrix entries are the factor correlation.

Structural model estimation and hypotheses testing

Descriptive statistics

The means and standard deviations for all constructs were determined and were displayed in Table 5. The highest mean of 3.56 was for the trialability, while the lowest mean for complexity was 2.30 on a scale of 1 to 5. The means for PU, PEU and behavioral intention were 3.79, 3.73, and 3.62, respectively.

Table 5: Descriptive statistics

| Construct (# Items) | Mean | Standard deviation |
|---------------------|------|--------------------|
| BI (six items) | 3.62 | .774 |
| PU (five items) | 3.79 | .708 |
| PEU (four items) | 3.73 | .709 |
| CPA (four items) | 3.54 | .808 |
| CPL (three items) | 2.30 | .769 |
| RA (five items) | 3.46 | .794 |
| OB (three items) | 3.39 | .930 |
| TRI (three items) | 3.56 | .794 |

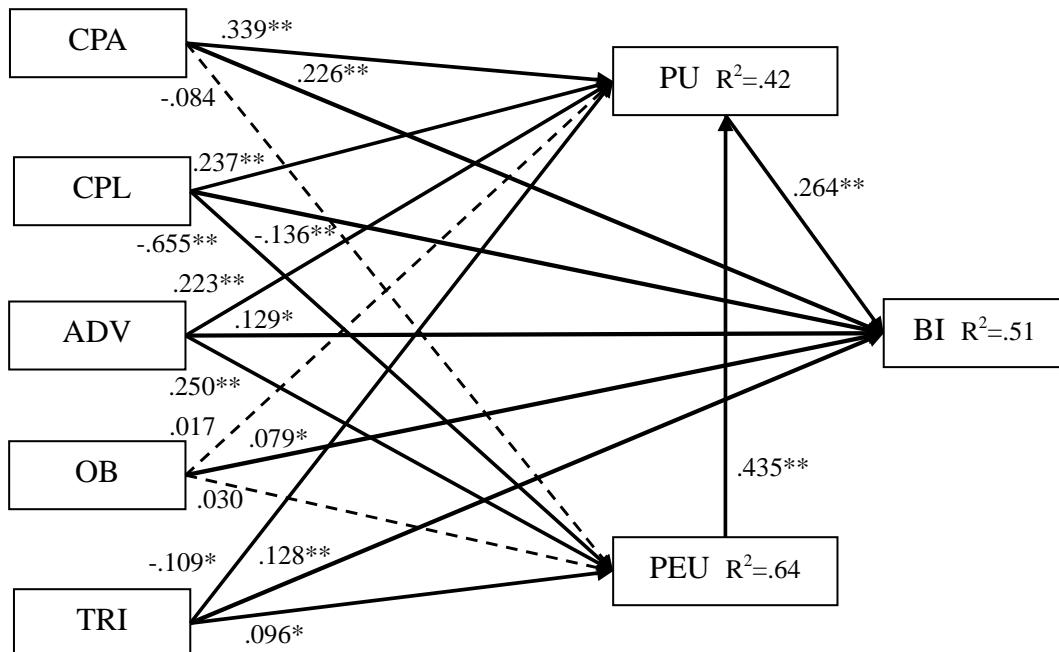
Structural equation modeling (SEM)

SEM was performed to test the fit between the research model (Figure 1) and the obtained data. This technique was chosen for its ability to simultaneously examine a series of dependence relationships, especially when there were direct and indirect effects among the constructs within the model (Hair, et al., 2006). The first step in interpreting SEM results includes reviewing fit indices, which provide evidence on how well the fit is between the data and the proposed structural model. If the model fits the data well enough, a second step involves reviewing the feasibility of each path in the model by examining whether the weights are statistically significant and practically significant. Practical significance is evaluated on the basis of whether the effect size estimation (the R^2) regarding a given path in the models is large enough.

In this study, Amos 6.0 was employed and the SEM estimation procedure was a maximum likelihood estimation. A similar set of fit indices was used to examine the structural model. Comparison of all fit indices with their corresponding recommended values provided evidence of a good model fit ($\chi^2/df = 1.42$, GFI = 0.95, AGFI = 0.93, CFI = 0.99, RMR = 0.02, and RMSEA = 0.03). The next step in the data analysis was to examine the significance and strength of hypothesized relationships in the research model. The results of the analysis of the structural model, including path coefficients, path significances, and variance explained (R^2 values) for each dependent variable presented in Figure 2.

Figure 2 showed the resulting path coefficients of the proposed research model. Overall, fourteen out of seventeen hypotheses were supported by the data. Three endogenous variables were tested in the model. The results showed that PU significantly influenced BI ($\beta = 0.267$, $P < 0.001$), supporting hypothesis H6-2. PEU was found to be significant in influencing PU ($\beta = 0.38$, $P < 0.001$), supporting hypotheses H6-1. Furthermore, PU was significantly influenced by two exogenous factors: CPA ($\beta = 0.33$, $P < 0.001$) and ADV ($\beta = 0.25$, $P < 0.001$) which support hypotheses H1-1 and H3-1. CPL ($\beta = 0.20$, $P < 0.01$) and TRI ($\beta = -0.099$, $P < 0.01$) which also significantly influenced PU. However, the effect was in contrast to what was hypothesized (H2-1 and H5-1). PEU was found to be significantly influenced by three exogenous factors: CPL ($\beta = -0.64$, $P < 0.001$), ADV ($\beta = 0.21$, $P < 0.001$), and TRI ($\beta = 0.09$, $P < 0.001$), supporting hypotheses H2-2, H3-2, and H5-2, respectively. Thus H1-2 and H4-2 were not supported. BI was significant influenced by all five exogenous variables: CPA ($\beta = 0.23$, $P < 0.001$), CPL ($\beta = 0.13$, $P < 0.001$), ADV ($\beta = 0.13$, $P < 0.05$), OB ($\beta = 0.08$, $P < 0.05$), and TRI ($\beta = 0.13$, $P < 0.05$). Therefore, hypotheses H1-3, H2-3, H3-3, H4-3, and H5-3 were supported.

PU was found to be significantly determined by six variables (CPA, CPL, ADV, OB, TRI, and PEU), resulting in an R^2 of 0.42. This meant that the above variables accounted for 42% of variance in PU. Likewise, PEU was found to be significantly determined by five exogenous variables (CPA, CPL, ADV, OB, and TRI), resulting in an R^2 of 0.64. This meant that the above exogenous variables explained for 64% of variance in PEU. BI was significantly determined by CPA, CPL, ADV, OB, TRI, and PU, resulting in an $R^2 = 0.51$. In other words, the variables described above explained 51% of the variance of BI. A summary of the hypotheses testing results is shown in Table 6.



(Note. * $p < .05$ ** $p < .01$)

Figure 2. Path test of the research model

Table 6: Hypotheses testing results

| Hypotheses | Path | Direction | Results |
|------------|-----------|-----------|---------------|
| H1-1 | CPA → PU | Positive | Supported |
| H1-2 | CPA → PEU | Negative | Not supported |
| H1-3 | CPA → BI | Positive | Supported |
| H2-1 | CPL → PU | Positive | Supported |
| H2-2 | CPL → PEU | Negative | Supported |
| H2-3 | CPL → BI | Negative | Supported |
| H3-1 | ADV → PU | Positive | Supported |
| H3-2 | ADV → PEU | Positive | Supported |
| H3-3 | ADV → BI | Positive | Supported |
| H4-1 | OB → PU | Positive | Not supported |
| H4-2 | OB → PEU | Positive | Not supported |
| H4-3 | OB → BI | Positive | Supported |
| H5-1 | TRI → PU | Negative | Supported |
| H5-2 | TRI → PEU | Positive | Supported |
| H5-3 | TRI → BI | Positive | Supported |
| H6-1 | PEU → PU | Positive | Supported |
| H6-2 | PU → BI | Positive | Supported |

Discussion

The research purpose was to develop a new hybrid technology acceptance model by combining TAM with IDT to explore the factors affecting employees' behavioral intentions to use the e-learning systems in organizations. This study was a pioneering effort in applying IDT into a TAM model. Based on our proposed model, we explored the relationships among five innovative characteristics with the PEU, the PU, and intention to use and the relationship between usefulness with employees' intention to use the e-learning systems. Overall, the results confirmed the research model and the hypotheses.

The results were consistent with previous studies showing that compatibility and relative advantages had significant positive effects on PU. It could be implied that prior to the employees' decision to use the e-learning systems, they tended to evaluate whether the e-learning systems could meet their job needs or be relevant to their job. If they perceived that the e-learning systems could meet their job needs, then they were likely to consider the e-learning systems to be useful. On the other hand, when the users regarded the e-learning systems as being better than the traditional training techniques or other approaches, they may perceive the e-learning systems to be more useful (Chang & Tung, 2008; Wu & Wang, 2005).

Secondly, contrary to the previous studies (Lee, 2007; Hardgrave, et al., 2003), our results showed that complexity had a significant positive effect on PU. We argued that when the e-learning systems were perceived to be of higher complexity, the employees tended to perceive higher usefulness of the e-learning systems. Although the employees felt great difficulty in operating the systems, they believed that e-learning systems helped improve their job performances. For those employees who considered e-learning systems to be simple, easy to understand or operate, they may not necessarily regard e-learning systems to be helpful to promote their job performance. Given these conflicting results, more research remained to be done to explore the relationship between these two constructs.

This study results also indicated that observability had no significant effect on PU, which was inconsistent with prior studies (Huang, 2004; Yang, 2007). The possible reason was that although the e-learning users could observe other individuals' use of e-learning systems and had an impression of how to operate them, they perceived the e-learning systems to be less useful in facilitating their job performances. Further rigorous follow-up studies should be conducted to investigate the precise role of observability in this area.

Moreover, the results suggested that trialability had a significant negative effect on PU. Specifically, the higher the trialability, the lower the PU would be. However, the test results were incongruent with the previous findings.

Our results strongly supported the hypothesis that complexity had a significant negative effect on PEU. In addition, the relative advantages and trialability had significant positive impacts on PEU. These results were consistent with previous research findings (Yang, 2007; Hardgrave et al., 2003). We uncovered that the employees would perceive the e-learning systems with unease, when they felt it was complex and difficult to operate. It implied that if the employees believed that the e-learning system could promote their job performances, they tended to think e-learning systems were easy to use. Additionally, when the employees had more opportunities to try the e-learning systems, they were more likely to view them as being easier to use. Hence, to enhance the perception of ease of use, the e-learning system developers should design the systems to be user-friendly and relevant to the employees' jobs. The managers should provide employees with organizational support for trying out e-learning systems.

Compatibility and observability had no significant effects on PEU and was consistent with previous research (Lin, 2006). This was inconsistent with the results from Huang (2004) and Yang (2007). Due to the mixed results, further investigation should individually examine the effects of specific innovative characteristics on PEU in various contexts of e-learning systems.

This study found that compatibility, complexity, relative advantages, observability, and trialability had significant effects on the employees' behavioral intention of using e-learning systems. These findings supported existing research that there existed strong relationships among the five innovative characteristics and the behavioral intention (Chang & Tung, 2008; Wu & Wang, 2005; Hardgrave et al., 2003). In order to promote the employees' intention to use the e-learning systems, e-learning systems designers should pay attention to the development of innovative characteristics and content of e-learning systems for potential users.

As prior research demonstrated, we found that the TAM appeared to provide researchers a theoretically sound and parsimonious model which can be used to predict the employees' behavioral intention to use the e-learning systems. According to TAM, PU and PEU had a significant positive and direct effect on the behavioral intention (BI) to use the e-learning systems. Such was the case in this study; the e-learning systems users thought that the higher PU resulted in a higher behavioral intention to use the e-learning systems. This study also indicated that PEU had a positive direct effect on PU. These findings echoed what Venkatesh & Davis, 2000) found in their study.

Furthermore, these findings supported existing research that found a strong relationship between PEU and PU, as originally proposed by Davis et al. (1989). If the e-learning systems are perceived as easy to use, employees may perceive them as useful. With the agreement of previous studies, both usefulness and ease of use were believed to be important factors in determining the acceptance of e-learning systems.

Limitations, suggestions and implications

Despite the careful attention to research methodology, improvements can be made in future studies in the following areas. Firstly, although the findings provided meaningful insights for the use of e-learning systems in organizational contexts, there may be a potential research bias in the sampling method due to the selection of a sample of willing respondents. To compensate for this drawback, future research should be conducted to test the proposed model using a random sampling approach. Furthermore, the study data were collected from the self-reported instrument. In fact, there could be a difference between what the participants responded to and what they actually did in terms of actual use of the e-learning systems. Hence, other methods of data collection (i.e. interviews and focus groups) and appropriate qualitative analyses should be conducted to provide a holistic understanding of the results of the current study. Next, this study was cross-sectional and not longitudinal. Therefore, it was uncertain whether the e-learning systems acceptance and usage behaviors were influenced by the individuals' expectations. Additionally, an individual's perceptions change over time when they gain more experience (Venkatesh & Davis, 2000; Venkatesh et al., 2003). Therefore, longitudinal research should be conducted to evaluate the validity of the proposed model and our findings.

The importance of the five innovative characteristics in affecting behavioral intention had several implications for researchers and practitioners. Firstly, TAM can be used as a cost-effective measurement to effectively predict the future use of e-learning systems. Secondly, according to the innovation diffusion theory (Rogers, 1995), adoption is not a snapshot and one-time decision, but rather a continuously staged process that can be investigated and boosted (Leonard-Barton, 1988). Potential users must first learn about the innovative tool and be persuaded to try it out

before they decide whether to adopt it. Therefore, this study suggested that well-designed trainings should be provided for the employees to familiarize themselves with the fundamental knowledge about how to use the e-learning systems as well as the trial opportunities to build a better understanding in the operational functions. The trainers' frequent demonstration of the use of e-learning systems help the employees form positive beliefs and attitudes, which in turn influences their behavioral intention and actual use of e-learning systems. As such, trainers should introduce and describe the benefits of e-learning systems and their relevance to their job performances. Additionally, trainers and system designers of e-learning systems should carefully consider the needs of e-learning system users and ensure that the e-learning systems effectively meet their job needs and demands.

Conclusions

This study has validated TAM and IDT in the organizational context and provided a further understanding into the employees' possible perceptions about the use of e-learning systems. The contribution of this study to e-learning acceptance research was discussed. While the merits of the TAM were manifested, the findings of this study provided greater insights when analyzing users' acceptance and adoption of e-learning systems. As a result, we proposed that the combination of TAM and IDT models could offer better overall results.

Furthermore, our findings suggested that system developers, designers, and institutional purchasers of e-learning systems carefully consider the needs of employees and ensure that selected systems effectively meet these demands. The five innovative characteristics with e-learning systems could be important determinants of user adoption of e-learning systems.

Reference

- Agarwal, R. (2000). Individual Acceptance of Information Technologies. *Educational Technology Research and Development*, 40, 90-102.
- Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences*, 30, 361-391.
- Agarwal, R., Sambamurthy, V., & Stair, R. M. (2000). The evolving relationship between general and specific computer efficacy: An empirical assessment. *Information Systems Research*, 11(4), 418-430.
- Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social behavior*, Englewood Cliffs, NJ: Prentice Hall.
- Arbaugh, J. B. (2002). Managing the on-line classroom: a study of technological and behavioral characteristics of web-based MBA courses. *Journal of High Technology Management Research*, 13, 203-223.
- Arbaugh, J. B., & Duray, R. (2002). Technological and structural characteristics, student learning and satisfaction with web-based courses: An exploratory study of two on-line MBA programs. *Management Learning*, 33(3), 331-347.
- Carter, L., & Be'langer, F. (2005). The utilization of e-government services: Citizen trust, innovation and acceptance factors. *Information Systems Journal*, 15(1), 5-25.
- Chan, T. C. (2005). *A Study of the Adoption e-Learning Factors on Taiwan*, M.S. Thesis, Chaoyang University of Technology.
- Chang, S. C., & Tung, F. C. (2008). An empirical investigation of students' behavioural intentions to use the online learning course websites. *British Journal of Educational Technology*, 39(1), 71-83.
- Chau, Y. K., & Hu, J. H. (2001). Information Technology Acceptance by Individual Professionals: A Model Comparison Approach. *Decision Sciences*, 32(4), 669-719.
- Chen, L. D., Gillenson, M. L., & Sherrell, D. L. (2002). Enticing online consumers: An extended technology acceptance perspective. *Information & Management*, 39(8), 705-719.
- Chin, W. C., & Todd, P. A. (1995). On the use, usefulness and ease of use of structural equation modeling in MIS research: a note of caution. *MIS Quarterly*, 19(2), 237-246.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical model. *Management Science*, 35(8), 982-1003.
- DeRouin, R. E., Fritzsche, B. A., & Salas, E. (2005). E-learning in organizations. *Journal of Management*, 31(6), 920-940.

- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 48, 39-50.
- Gefen, D. (2004). What makes an ERP implementation relationship worthwhile: Linking trust mechanisms and ERP usefulness. *Journal of Management Information*, 21(1), 263-288.
- Hair, J. F., Jr, Black, B., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Hardgrave, B., Davis, F., & Riemenschneider, C. (2003). Investigating Determinants of Software Developers' Intentions to Follow Methodologies. *Journal of Management Information Systems*, 20(1), 123-151.
- Horton, R., Buck, T., Waterson, P., & Clegg, C. (2001). Explaining intranet use with the technology acceptance model. *Journal of Information Technology*, 16, 237-249.
- Huang, L. Y. (2004). *A study about the key factors affecting users to accept Chunghwa Telecom's Multimedia on Demand*. Unpublished Master Thesis, National Sun Yat-Sen University.
- Igbaria, M., Zinatelli, N., Cragg, P., & Cavaye, A. L. M. (1997). Personal computing acceptance factors in small firms: A structural equation model. *MIS Quarterly*, 21(3), 279-305.
- Johansen, R., & Swigart, R. (1996). *Upsizing the Individual in the Downsizing Organization: Managing in the Wake of Reengineering, Globalization, and Overwhelming Technological Change*. Reading, MA: Addison-Wesley.
- Karahanna, E., Straub, D.W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly*, 23, 183-213.
- Liao, C., Palvia, P., & Chen, J. (2009). Information technology adoption behavior life cycle: Toward a Technology Continuance Theory (TCT). *International Journal of Information Management*, 29(4), 309-320.
- Lau, S. H., & Woods, P. C. (2008). An Investigation of User Perceptions and Attitudes toward Learning Objects. *British Journal of Educational Technology*, 39(4) 685-699.
- Lee, Y. H. (2007). Exploring key factors that affect consumers to adopt e-reading services. Unpublished Master Thesis, Huafan University.
- Legris, P., Ingham, J. & Colerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information and Management*, 40, 191-204.
- Ong, C. S., Lai, J. Y., & Wang, Y. S. (2004). Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies. *Information & Management*, 41, 795-804.
- McFarland, D. J., & Hamilton, D. (2006). Adding contextual specificity to the technology acceptance model. *Computers in Human Behavior*, 22, 427-447.
- Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3), 192-222.
- Mun, Y. Y., & Hwang, Y. (2003). Predicting the use of Web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and TAM. *Human-Computer Studies*, 59(4), 431-449.
- Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-based Virtual Learning Environments: A Research Framework and a Prdimentary Assessment of Effectiveness in Basic Skills IT training. *MIS Quarterly*, 25(4), 401-426.
- Pituch, K. A., & Lee, Y. K. (2006). The influence of system characteristics on e-learning use. *Computer & Education*, 47, 222-244.
- Rogers, E. M. (1995). *Diffusion of innovations* (4th Ed.). New York: Free Press.
- Segars, A. H., & Grover, V. (1993). Re-examining PEU and usefulness: a confirmatory factor analysis. *MIS Quarterly*, 17(4), 517-525.
- Shih, C. H. (2007). *Integrating Innovation Diffusion Theory and UTAUT to Explore The Influencing Factors on Teacher Adopt E-learning System – With MOODLE as an Example*. Unpublished Master Thesis, Dayeh University.
- Sigala, M., Airey, D., Jones, P., & Lockwood, A. (2000). The diffusion and application of multimedia technologies in the tourism and hospitality industries. *Information and Communication Technologies in Tourism 2000*, Wien: Springer, 396-407.
- Taylor, S., & Todd, P. A. (1995). Understanding information technology usage: a test of competing models. *Information Systems Research*, 6(1), 144-176.

- Taylor, S., & Todd, P. A. (1995). Assessing IT usage: The role of prior experience. *Management Information Systems Quarterly*, 19(4), 561-570.
- Venkatesh, V., & Davis F. D., 2000. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Management Science*, 46(2), 186-204.
- Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.
- Xu, Z., & Yuan, Y. (2009). Principle-based dispute resolution for consumer protection. *Knowledge-Based Systems*, 22(1), 18-27.
- Wu, J. H., & Wang, S. C. (2005). What drives mobile commerce? An empirical evaluation of the revised technology acceptance model. *Information Management*, 42, 719–729.
- Wu, J. P., Tsai, R. J., Chen, C. C., & Wu, Y. C. (2006). An integrative model to predict the continuance use of electronic learning systems: Hints for teaching. *International Journal on E-Learning*, 5(2), 287-302.
- Yang, M. M. (2007). *An exploratory study on consumers' behavioral intention of usage of third generation mobile value-added services*, Unpublished Master Thesis, National Cheng Kung University.

Using Wikis for Learning and Knowledge Building: Results of an Experimental Study

Joachim Kimmerle*, Johannes Moskaliuk and Ulrike Cress¹

University of Tuebingen, Department of Applied Cognitive Psychology and Media Psychology, Germany //

¹Knowledge Media Research Center, Germany // j.kimmerle@iwm-kmrc.de // j.moskaliuk@iwm-kmrc.de // u.cress@iwm-kmrc.de

*Corresponding author

ABSTRACT

Computer-supported learning and knowledge building play an increasing role in online collaboration. This paper outlines some theories concerning the interplay between individual processes of learning and collaborative processes of knowledge building. In particular, it describes the co-evolution model that attempts to examine processes of learning and knowledge building by working on wikis. We report an experimental study that aimed at testing some predictions of this model empirically. The results support the assumption that accommodative knowledge building and a development of conceptual knowledge takes place particularly when there is incongruity at a medium level between people's knowledge and the information contained in a digital artefact. In contrast, assimilative knowledge building and the development of factual knowledge depends largely on people's prior knowledge. Concluding, the consequences of these findings on educational uses of wikis are discussed.

Keywords

Learning, Knowledge building, Wiki, Experiment

Introduction

In the last few years, various technologies and tools have been developed that provide opportunities for computer-supported knowledge exchange (Cress, Kimmerle, & Hesse, 2006; Kimmerle, Cress, & Hesse, 2007), as well as for computer-assisted learning and collaborative knowledge building (Bryant, 2006). Technologies that support people in communicating, interacting, and collaborating in large communities are referred to as "social software" (cf. Kolbitsch & Maurer, 2006). The collaborative development of knowledge that is enabled by social software illustrates quite nicely what Scardamalia and Bereiter (1994, 1996, 2003) describe in their theory of knowledge building. Knowledge building in the web is virtually scaled up from smaller groups (e.g., a class or a team in an organization) to large communities of users. *Wikipedia*, the Online Encyclopaedia, is one of the most frequently cited examples of this phenomenon (cf. Baytiyeh & Pfaffman, 2010; Goldspink, 2010): 13.8 million registered and many other unregistered users have contributed more than 3.5 million articles to the English version of the Wikipedia (figures of January 2011). This success story of social software on the Internet has convinced many people to apply social-software tools in educational contexts as well (Evans, 2008; Fessakis, Tatsis, & Dimitracopoulou, 2008; Kim, 2008). It is assumed that social software has a great potential in the context of learning and knowledge building, both in formal and informal learning situations (Bryant, 2006; Parker & Chao, 2007; Wang & Turner, 2004).

Wikis are particularly interesting for learning purposes (Reinhold, 2006; Shih, Tseng, & Yang, 2008; Wang & Turner, 2004; Yukawa, 2006). Wikis may be characterized as collections of websites on intranets or the Internet. Such websites cannot only be read by users, but may also be edited by any participant (Leuf & Cunningham, 2001). In a wiki, people may generate content and link it to other content, using hyperlinks. Users are allowed to change text, insert new text, or even delete the whole text of a wiki or parts of it. In this way, a community of wiki users can work together in order to create one shared digital artefact. Thus, working on a wiki enables a collaborative development of knowledge (Köhler & Fuchs-Kittowski, 2005) and leads to knowledge building as the creation of new and innovative knowledge. Users can use a wiki to share their knowledge, create a joint artefact, discuss and integrate different opinions, develop innovative ideas. This may, at the same time, lead to individual learning. Thus, wikis may be considered as powerful tools for learning and knowledge building in educational contexts.

The initial enthusiasm has settled down, however. Some practitioners report on the pitfalls of social software in educational contexts (cf. Cole, 2009). Often, students use wikis as a read-only-source. They will not change any content and seem not to be motivated to construct knowledge together with others. It seems that the success of a real social web (e.g., Wikipedia) cannot directly be transferred into classroom (Forte & Bruckman, 2006). Using wikis in

educational contexts will not *per se* lead to learning and knowledge building. So the main issue of this paper is to examine the process of individual learning and collaborative knowledge building with wikis and specify the circumstances under which these processes may be successful. The first part of the paper gives a short survey of psychological and educational approaches that are helpful to understand individual learning and collaborative knowledge building with wikis. In the second part of the paper, we report on an experimental study that focused on incongruity between the knowledge of individuals and the information contained in a wiki, as an important trigger for processes of learning and knowledge building. The final part of the paper aims at integrating theoretical assumptions and empirical results, and suggests further implication for the use of wikis and other social software in educational contexts.

Individual learning and collaborative knowledge building

When looking at theories and research about computer-supported collaborative learning and knowledge building (cf. Tsatsos, Andreas, & Pomportsis, 2010), it is noticeable that we often find two categories of theoretical approaches: on the one hand, theories that primarily deal with individual results of learning processes, and, on the other hand, theories that are mainly interested in collective processes. What is needed, however, to really understand learning and knowledge building with social software is theories or models that look in more depth into the *interplay* between the individual and group level of learning. We will now present some of these approaches that aim at examining both levels and their interplay.

Knowledge building as a theoretical term (Hewitt & Scardamalia, 1998; Scardamalia, 2002; Scardamalia & Bereiter, 2003) describes the creation of new knowledge and innovative ideas as a socio-cultural process that takes place within a community. It aims at creating "... something of value to the community – theories, explanations, problem formulations, interpretations, and so on, which become public property that is helpful in understanding the world ..." (Scardamalia & Bereiter, 1999, p. 276). The term knowledge building refers to the collective creation of public knowledge. This *collective* effort is in the main focus of that theory. This approach is aware that a collective is composed of individuals, but the main focus is on what occurs within the community. The group is considered as more than an assemblage of individual people. From this perspective, individual learning is always considered as well, but the main interest is not directed onto this process. Individual learning is, so to speak, rather a by-product of the constructive process of collective knowledge building.

The *knowledge creation* model (Nonaka, 1994; Nonaka & Takeuchi, 1995; Nonaka & Toyama, 2003) considers knowledge dissemination as the aim of any knowledge organization. In this model, however, most of the knowledge that exists in an organization is part of the experience of individuals; it is tacit knowledge (Polanyi, 1966). This tacit knowledge can hardly be expressed verbally and transferred to others without own experience. The authors describe a dynamic process of socialisation, externalization, combination, and internalization that is necessary to share tacit knowledge.

Another approach that tries to combine the perspectives of individual learning and collaborative knowledge building is *group cognition* (Stahl, 2006, 2009). Its focus is on interaction between individuals and the group. In this theory, learning is constituted in a context of interaction with others, and small groups of learner are considered as the engines of knowledge construction. A shared digital artefact serves as mediator between individual knowledge and knowledge of the group, and it is a tool for knowledge construction at the same time.

Activity theory (Engeström, 1987, 1990, 2001) puts learning in a cultural context. Individuals use cultural artefacts such as tools, symbols, or language to solve problems and fulfil tasks. Activity theory is interested in cognitive competences of individuals, but it assumes that it does not make much sense to consider individuals as isolated entities, because human cognition and learning are always part of a society or community, based on interaction between individuals and some (cultural and social) context. Therefore, human cognition should always be considered as socially embedded and its analysis needs to take this aspect into account. Various other theories have focused on this aspect of cultural, situational, or physical context of individual learning, e.g., the *distributed cognition* approach (Hutchins, 1995a, 1995b), the *situated cognition* approach (Greeno, 1998), and the concept of *embodied cognition* (Clark, 1997; Gibbs, 2006).

What becomes clear from this short survey is that it is widely accepted that we have to consider three dimensions in this context: cognitive processes of individuals, social processes inside a community, and interaction between individuals and a community. The theories presented here are very heterogeneous, and they were originally not developed for research on wikis. Research on learning and knowledge building in the social web requires a theoretical framework that considers these seminal theories, but that focuses on the tight conjunction between individual learning and collaborative knowledge building and that permits, at the same time, the formulation of concrete hypotheses about conditions for successful learning and knowledge building with wikis. In the following, we will present a model that is, in our opinion, suitable for that purpose.

The co-evolution model

The co-evolution model by Cress and Kimmerle (2007, 2008; Kimmerle, Cress, & Held, 2010) aims at describing and explaining how individual learning and collaborative knowledge building with social software – with wikis in particular – may take place (cf. also Kimmerle, Moskaliuk, Cress, & Thiel, 2011). This model adopts a systems-theoretical perspective by referring to the work of Luhmann (1990, 2006). It differentiates between the cognitive systems of users and the social system established by the wiki community. These systems differ in their mode of operation: whereas cognitive systems are based on cognitive processes (thinking, reasoning, consciousness), social systems operate via communication (Luhmann, 1995). In order to describe processes of individual learning in the cognitive systems of individuals, Cress and Kimmerle borrowed some ideas from Piaget's model of equilibration (Piaget, 1970, 1977). In addition, they take up Piaget's considerations and translate them into processes that occur within social systems in order to describe collaborative knowledge building.

Piaget's model of equilibration (1985) assumes that people take up new information from the environment and integrate it into their existing knowledge. If new information is not in line with prior knowledge, this will lead to cognitive conflicts. Individuals have two options to deal with such cognitive conflicts: in simple terms, they may either assimilate new information or accommodate their existing knowledge. In the process of assimilation, people utilize their prior knowledge, as it is, in order to understand the new information. Accommodation, however, is a process in which people modify their existing knowledge as a result of handling new information. So in this accommodation process, individuals will qualitatively change their prior knowledge, not only assimilate additional information into their existing knowledge.

Cress and Kimmerle (2007, 2008) suppose that accommodation and assimilation will not only occur in cognitive systems, but also in social systems. In the case of wikis, assimilation is characterized by purely adding information, without explicitly connecting it to any previously existing information. The organization of the wiki remains unaffected; the wiki simply “assimilates” this new information. Accommodation is characterised by re-arranging the existing information in that wiki. This is the case, for example, when sections are rewritten, the arrangement of pages is re-organized, or previously existing and new information are combined and incorporated into each other (cf. also Majchrzak, Wagner, & Yates, 2006). This leads to four processes of individual learning and collaborative knowledge building: assimilative learning, accommodative learning, assimilative knowledge building and accommodative knowledge building.

This model integrates individual and collaborative aspects. It describes individual learning as internalization of information from the social system into the cognitive system, and collaborative knowledge building as externalization of individual knowledge from the cognitive system into the social system. These two systems are not independent, but influence each other (Kimmerle, Moskaliuk, Harrer, & Cress, 2010). Their development is a mutual progress, a *co-evolution*. Besides, the co-evolution model specifies the conditions under which individual learning and collaborative knowledge building may take place. The authors states that incongruities between people's knowledge and the information contained in a wiki can lead to cognitive conflicts in the sense of Piaget. They specify their assumption by saying that the *extent* of incongruities has an impact on the extent of the cognitive conflict, which in turn affects learning and knowledge building. Incongruities *at a medium level* are supposed to be particularly conducive to both individual learning and collaborative knowledge building.

The benefit of the co-evolution model is the fact that it offers an integrative point of view and considers cognitive and social systems at the same time. Apart from that, it specifies the configuration of both systems that leads to an

increase of individual learning and collaborative knowledge building at the same time. The experimental study that we present in the following section will examine these aspects and test the predictions of the model.

Experimental study

Moskaliuk, Kimmerle, and Cress (2009) were able to demonstrate in their study that incongruities at a medium level do indeed support individual learning and collaborative knowledge building to a higher extent than high or low incongruities. This study worked with participants who were supplied with a high level of prior knowledge, and the authors operationalized incongruities by making their participants work on wikis that were either completely filled with information (low incongruity), half-filled with information (medium incongruity), or completely empty (high incongruity).

Concerning assimilative knowledge building, Moskaliuk et al. (2009) found an equally high level in the medium- and the high-incongruity condition. This is not surprising, because in both conditions participants had enough room to write things down. In the low-incongruity condition, however, with a virtually full wiki, they did not have this opportunity. Despite the fact that participants in the medium-incongruity condition did not add more words to the wiki (assimilative knowledge building) than in the high-incongruity condition, there was still more accommodative knowledge building, and participants also learned more (developed more factual and more conceptual knowledge). The authors concluded that simply writing down many words will not necessarily lead to more learning. What is more important in this context is the opportunity to link one's own knowledge to existing information (in the wiki), which is not the case in the low- and the high-incongruity conditions. In order to be able to link one's own knowledge to existing information, there has to be a certain degree of agreement between knowledge (in the individuals' minds) and information (in the wiki), but this agreement should not be exhaustive. In the study by Moskaliuk et al. (2009), participants had the opportunity to externalize their own knowledge (except for the low-incongruity condition). The results of this experiment suggest that some opportunities should exist to externalize one's own knowledge, and that this is an important precondition for learning and knowledge building.

The question arises, however, what will happen in such a knowledge-building situation if this opportunity for externalizing one's own knowledge is extremely reduced? We assume that people will then still develop conceptual knowledge and participate in accommodative knowledge building – as long as there is a sufficient “overlap” between individual and collective knowledge, as it is the case with medium incongruity. This is because in such a situation people have the opportunity to elaborate on their prior knowledge, as it will be stimulated and inspired by novel input. These processes are supposed to lead to improved understanding. The same is true for information in the wiki, which is also challenged and stimulated in this process, and this, in turn, may also lead to qualitative improvement.

Assimilative processes, however, are supposed to be primarily depending on people's individual prior knowledge: if they only have little prior knowledge, they will hardly be able to add new information to a wiki (assimilative knowledge building), and if they know a great deal in advance, they will still possess more factual knowledge afterwards. These considerations lead to concrete hypotheses that will be presented in detail in the next paragraph. Subsequently, we will describe how we managed to remove the opportunity to externalize one's own knowledge in the study reported here.

Hypotheses

In our opinion, the superiority of a medium degree of incongruity between people's knowledge and information contained in a wiki refers particularly to accommodative processes. The impact of incongruity on assimilative processes, however, is supposed to depend on the respective operationalization of incongruity. In order to examine these specifying assumptions and come to more differentiated conclusions, we conducted a pivotal modification compared to the study by Moskaliuk et al. (2009): we decided not to keep the individuals' prior knowledge constant (it was kept at a constantly high level in that study) and vary the information in the wiki, but instead operationalize incongruity the other way round: information in the wiki was kept constant (at a high level) and people's prior knowledge varied (Figure 1).

| low incongruity | | medium incongruity | | | | high incongruity | |
|-----------------|-------|--------------------|-------|------|-------|------------------|-------|
| wiki | part. | wiki | part. | wiki | part. | wiki | part. |
| B1 | B1 | B1 | | B1 | B1 | B1 | |
| B2 | B2 | B2 | | B2 | B2 | B2 | |
| B3 | B3 | B3 | | B3 | B3 | B3 | |
| B4 | B4 | B4 | | B4 | B4 | B4 | |
| S1 | S1 | S1 | S1 | S1 | | S1 | |
| S2 | S2 | S2 | S2 | S2 | | S2 | |
| S3 | S3 | S3 | S3 | S3 | | S3 | |
| S4 | S4 | S4 | S4 | S4 | | S4 | |

Figure 1. Distribution of information provided in the wiki and participants' prior knowledge ("part.") in three experimental conditions. B1 to B4 are arguments that describe biological causes of schizophrenia, S1 to S4 arguments that describe social causes

Hypothesis 1. We expect the lowest degree of assimilative knowledge building in the high-incongruity condition (if people have no prior knowledge, it will be very difficult for them to write down anything). In more concrete terms, we expect more assimilative knowledge building in the medium-incongruity condition than in the high-incongruity condition (*Hypothesis 1a*). And we expect more assimilative knowledge building in the low- than in the high-incongruity condition (*Hypothesis 1b*).

Hypothesis 2. We expect the highest degree of accommodative knowledge building in the medium-incongruity condition, because this situation is most suitable when the point is to dispute and challenge existing information, i.e. when it comes to processes which are supposed to lead to qualitative improvement of that information. In more concrete terms, we expect more accommodative knowledge building in the medium- than in the low-incongruity condition (*Hypothesis 2a*). And we expect more accommodative knowledge building in the medium- than in the high-incongruity condition (*Hypothesis 2b*).

Hypothesis 3. We expect that people will possess more factual knowledge, the higher their respective knowledge was in the first place (i.e. the more facts they were familiar with in advance). In more concrete terms, we expect people to develop more factual knowledge in the low- than in the medium-incongruity condition (*Hypothesis 3a*). And we expect people to develop more factual knowledge in the medium- than in the high-incongruity condition (*Hypothesis 3b*).

Hypothesis 4. We expect that people will develop more conceptual knowledge in the medium-incongruity condition, because this condition is supposed to be optimal for elaborating on prior knowledge and gaining new insights. In more concrete terms, we expect that people will develop more conceptual knowledge in the medium- than in the low-incongruity condition (*Hypothesis 4a*). And we expect that people will develop more conceptual knowledge in the medium- than in the high-incongruity condition (*Hypothesis 4b*).

Method

In order to test these hypotheses empirically we performed an experimental study under controlled laboratory conditions. In this section, we will describe the experimental setting, the procedure adopted, and the measures that were applied.

Experimental setting. The experiment was conducted in groups with five to ten participants. Going through the complete experiment took about two hours. The wiki was introduced to the participants as a real wiki about Clinical Psychology, currently under development in order to inform patients or other interested people. The participants were made to believe that the other participants had to deal with a different part of the same wiki, i.e., on a different topic (paranoia, depression...), but in fact, each participant worked, independently from the others, with the same simulation of a wiki page (with faked prior versions, faked previous authors, faked edit dates) on the "causes of schizophrenia". We used nearly the same experimental material as in the study of Moskaliuk et al. (2009): eight different arguments (four "social" and four "biological" arguments) were used to build the wiki page about causes of

schizophrenia – Moskaliuk et al. (2009) had used ten arguments; apart from social and biological arguments they had also used two integrative arguments. We decided to omit these integrative arguments, because we wanted participants to arrive at an integrative approach on their own, because we think that this is a very good indicator of accommodation (see below). The eight arguments were the same in all conditions, keeping the information in the wiki constant. In order to manipulate incongruity between prior knowledge and information in the wiki, participants in the different experimental conditions received a different number of short texts in the look-and-feel of scientific newsletters before they started working on the wiki. Each newsletter text contained one argument (either social or biological). These newsletters dealt with the same eight arguments that were also part of the wiki page, but they were complemented by additional information that was irrelevant to the significance of the argument.

Participants in the low-incongruity condition received all eight newsletters in order to guarantee high prior knowledge. Participants in the medium-incongruity condition received either four newsletters with social arguments on causes of schizophrenia or four newsletters with biological arguments (in order to take into account potential qualitative differences between the social and the biological arguments, we provided two versions of the medium-incongruity condition). Participants in the high-incongruity condition received no newsletters. Consequently, their prior knowledge was low. Figure 1 shows the three experimental conditions.

Procedure. We used mobile computers for presenting two questionnaires (at the beginning and end of the experiment), for the instructions, and for a short tutorial that introduced the handling of the wiki. Before the participants started working on the wiki, they received the newsletters (except in the high-incongruity condition). The time that participants could spend reading the newsletters (learning phase) was determined, depending on the condition that they had been assigned to (see below).

Measures. Corresponding to the four processes of learning and knowledge building, there were four different dependent variables. For measuring modifications in the wiki, we made a log-file analysis for each participant, to compare the initial version of the wiki page to the last version at the end of the experiment. To measure assimilative knowledge building, we counted the number of words that participants had added to the wiki text. In order to measure accommodative knowledge building, we developed an index as follows: we counted those phrases that participants either used to refer to interaction between social and biological causes of schizophrenia (integrative arguments) or which they used to connect arguments (e.g., “on the one hand … on the other hand”, “in contrast”, “however”).

To measure factual knowledge about causes of schizophrenia (in the post-experimental questionnaire), we used the multiple choice test from Moskaliuk et al. (2009) with statements about the causes of schizophrenia. To measure conceptual knowledge, we asked participants to provide the best argument to explain why schizophrenia occurs. Their answers were rated by experts to distinguish between different levels of conceptual knowledge. An answer that contained a simple explanation (biological or social) was rated with one point. Two points were assigned when participants named both biological and social causes. They received three points if they pointed out to some kind of interaction between social and biological factors. And participants received four points if they were able to explain that it is external stress (social) that can uncover an inherent (biological) vulnerability. The number of points attained was taken to represent the degree of conceptual knowledge.

Experimental design. The experiment represented a between-subjects design with incongruity as between-subjects factor with three levels (low vs. medium vs. high). To distinguish clearly between the conditions, we varied the time that participants were able to spend on the development of prior knowledge by reading the scientific newsletters. In the low-incongruity condition they had 20 minutes (for eight newsletters); in the medium-incongruity condition they had 10 minutes (for four newsletters). Participants in the high-incongruity condition were not supposed to have any prior knowledge, so they were not given any learning time before working on the wiki. Assimilative and accommodative knowledge building as well as development of factual and conceptual knowledge served as dependent variables.

Participants. This study was carried out with 72 participants. 55 of these were women, 17 men. Their mean age was 22.06 years ($SD=3.48$). The participants were students from a local university. Students of Psychology were not allowed to participate, because of their potential prior knowledge on schizophrenia. The participants were randomly assigned to one of the experimental conditions. 25 participants were assigned to the low-, 25 to the medium-, and 22 to the high-incongruity condition.

Results

In order to test the hypotheses, we conducted independent sample t-tests for comparing the experimental conditions for each dependent variable.

H1a assumed more assimilative knowledge building in the medium- than in the high-incongruity condition. The data supported this hypothesis: $t(44)=4.78$, $p<.01$, $d=1.46$. There was more assimilative knowledge building in the medium-incongruity condition: $M_{med}=84.00$ ($SD=64.92$) vs. $M_{high}=11.71$ ($SD=26.09$).

H1b assumed more assimilative knowledge building in the low- than in the high-incongruity condition. The data supported this hypothesis: $t(43)=5.21$, $p<.01$, $d=1.59$. There was more assimilative knowledge building in the low-incongruity condition: $M_{low}=89.00$ ($SD=63.38$) vs. $M_{high}=11.71$ ($SD=26.09$).

H2a assumed more accommodative knowledge building in the medium- than in the low-incongruity condition. The data supported this hypothesis: $t(47)=3.55$, $p<.01$, $d=1.02$. There was more accommodative knowledge building in the medium-incongruity condition: $M_{med}=2.04$ ($SD=1.79$) vs. $M_{low}=0.58$ ($SD=0.93$).

H2b assumed more accommodative knowledge building in the medium- than in the high-incongruity condition. This hypothesis was also supported: $t(44)=4.36$, $p<.01$, $d=1.34$. There was more accommodative knowledge building in the medium-incongruity condition: $M_{med}=2.04$ ($SD=1.79$) vs. $M_{high}=0.29$ ($SD=0.46$). Figure 2 presents the results of H2.

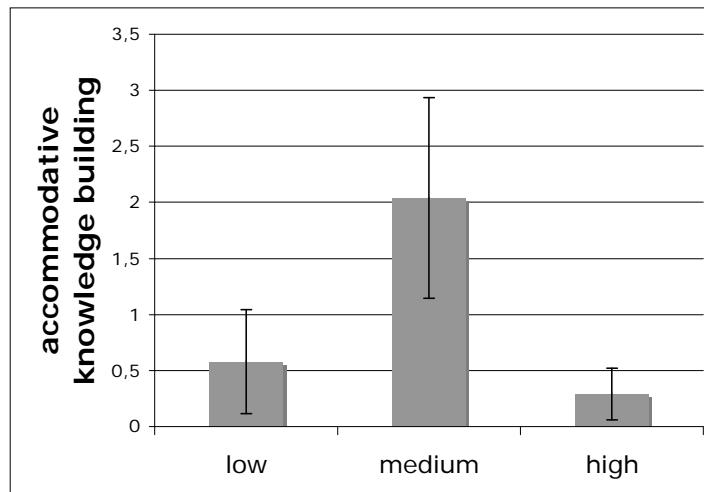


Figure 2. Accommodative knowledge building in the three conditions (means and standard deviations)

H3a assumed that people would possess more factual knowledge in the low- than in the medium-incongruity condition. The data supported this hypothesis: $t(48)=1.69$, $p=.05$, $d=0.48$. Participants in the low-incongruity condition had more factual knowledge: $M_{low}=14.72$ ($SD=2.01$) vs. $M_{med}=13.68$ ($SD=2.34$).

H3b assumed more factual knowledge in the medium- than in the high-incongruity condition. The data supported this hypothesis: $t(45)=2.31$, $p=.01$, $d=0.67$. Participants in the medium-incongruity condition possessed significantly more factual knowledge: $M_{high}=11.82$ ($SD=3.17$) vs. $M_{med}=13.68$ ($SD=2.34$).

H4a assumed that people would develop more conceptual knowledge in the medium- than in the low-incongruity condition. The data tended to support this hypothesis; this is, however, only a marginal effect: $t(48)=1.34$, $p=.09$, $d=0.38$. Participants in the medium-incongruity condition tended to develop more conceptual knowledge: $M_{low}=1.20$ ($SD=0.65$) vs. $M_{med}=1.48$ ($SD=0.82$).

H4b assumed more conceptual knowledge in the medium- than in the high-incongruity condition. The data supported this hypothesis: $t(45)=2.07$, $p=.02$, $d=0.61$. Participants in the medium-incongruity condition developed significantly more conceptual knowledge: $M_{high}=1.05$ ($SD=0.58$) vs. $M_{med}=1.48$ ($SD=0.82$). Figure 3 shows the results of H4.

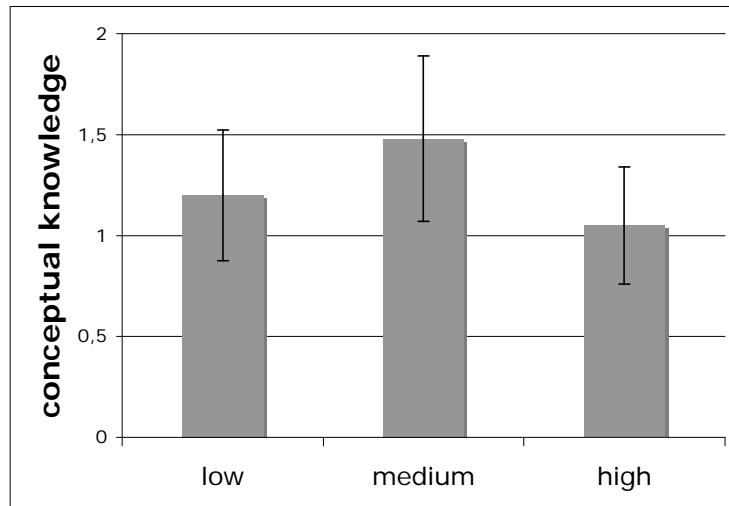


Figure 3. Conceptual knowledge in the three conditions (means and standard deviations)

Discussion

The results of this study show a higher extent of assimilative knowledge building in the low- and in the medium-incongruity condition than in the high-incongruity condition. Moreover, individuals possessed more factual knowledge in the low- than in the medium-incongruity condition and more factual knowledge in the medium- than in the high-incongruity condition. As expected, participants possessed, after the experiment, the more factual knowledge the higher their prior knowledge had been in the first place. These results are not too surprising and arise predominantly from the way incongruity was operationalized in this study: there is a simple linear relationship between the quantity of available learning material and the availability of factual knowledge.

But as far as accommodative processes are concerned, the results are rather revealing: in the medium-incongruity condition, more accommodative knowledge building occurred than in the other two conditions. And we found that people developed more conceptual knowledge in the medium-incongruity condition than in the low- and the high-incongruity condition (even though the difference between medium and low was only marginally significant). What is important here is the finding that people participated more in accommodative knowledge building and developed more conceptual knowledge in the case of medium incongruity; this cannot be explained with the length of learning time: allocating more learning time did not lead to a greater extent of accommodative knowledge building and development of conceptual knowledge. The relevant factor here is incongruity between information in the wiki and the individuals' knowledge, and the results indicate clearly that a medium level of this incongruity is particularly conducive to knowledge building and individual learning in terms of quality.

As compared to the study by Moskaliuk et al (2009), the present study aimed specifically at enabling processes of internalization of knowledge. In the high-incongruity condition, participants were able, while working on the wiki, to internalize all arguments from the wiki. In the medium-incongruity condition, it was only possible to internalize those parts of the arguments that had not previously been available in the newsletters. In the low-incongruity condition, participants found no new information in the wiki and would hardly have been able to develop any new knowledge by working on the wiki.

The experimental manipulation of opportunities for the internalization of knowledge had the consequence that opportunities provided for externalization were very low in all three conditions, because the wiki contained all the relevant information and participants could practically only add minor details. So this study is a mirror-inverted replica of the study by Moskaliuk et al. (2009), where opportunities for internalization were kept constantly low in all three conditions and those provided for externalization of knowledge were manipulated. Participants in this study had received all newsletters in all three conditions (low, medium, and high incongruity) and it was the different degrees of completeness of the wiki that made the difference between the three conditions. A comparison of the

results of the present study with the results from Moskaliuk et al. (2009) supports the assumption that it is mainly construction processes of externalization that act as catalysts of knowledge development.

To sum up, the experiment reported here support the assumptions of the model by Cress and Kimmerle (2008). Although in all experimental conditions users always had access to all information – it was only distributed differently between the newsletters and the wiki – the experimental conditions differed, in line with theory, in the extent to which they facilitated learning and knowledge building.

Conclusion

What are the implications of these findings for the use of wikis in educational contexts? Based on our own research and the theoretical approaches discussed in the first section of the paper, we can point out that an integrative point of view is essential to understand learning and knowledge building. We have to focus on the individual and individual learning processes on the one hand, as well as on the group or community and its knowledge-building processes on the other. In schools and universities often only individual achievement is considered. In other contexts (e.g., in enterprises or in a scientific community), however, knowledge is very often built jointly. So a broader understanding of learning and knowledge building needs both perspectives.

It is necessary to specify these processes of learning and knowledge building. The co-evolution model presented here describes accommodation and assimilation as two ways of integrating own knowledge and new information in a wiki. Besides, the model postulates that medium incongruity between one's own knowledge and new information leads to a greater cognitive conflict than low and high incongruity and to a greater extent of learning and knowledge building in terms of quality.

Concerning the use of wikis in educational contexts, we would like to make some practical suggestions (even though we are aware that a transfer of findings from laboratory studies to real world contexts will always need to be handled with care):

- One factor that triggers learning and knowledge building is the interaction between a person's individual knowledge and new information in the environment. A central catalyst here is a perceived cognitive conflict that may lead to accommodation and assimilation. So educators should allow cognitive conflicts. They should provide educational settings in which such conflicts may occur and may actively be solved by learners.
- So educators need to find an adequate level of incongruity. Starting with an empty wiki disables the possibility of linking one's own knowledge to available content, and a complete wiki with all relevant information may demotivate students from participating. A partly filled, but still incomplete wiki provides the opportunity to consider new information, but still add one's own knowledge, settle controversies, and integrate different positions.
- Educators need to specify their learning goals. Working with a wiki seems to have a lower impact on the development of factual knowledge and assimilative knowledge building than on the development of conceptual knowledge and accommodative knowledge building. If the aim is merely to impart facts, then wikis are not necessarily the first choice. But if the goal is to discuss and integrate different aspects of one topic, or to develop mutual understanding of complex issues, a wiki seems to be an appropriate tool.

A conclusion of the theoretical background and the study reported here is that wikis seem to be suitable instruments to encourage and facilitate processes of individual learning and collaborative knowledge building at the same time. Wikis are more than a piece of software technology – they stand for a constructivist view on learning, one that defines knowledge building as a socio-cognitive process between individuals and groups. So, this tool and the underlying principles (cf. Moskaliuk & Kimmerle, 2009) may be used on a broader scale, both in informal and institutionalized learning arrangements. The underlying theoretical paradigm, however, cannot only be applied to wikis. It may also be transferred to situations in which people deal with other types of shared digital artefacts. We conclude that it is an adequate approach to consider both the processes in the cognitive systems of individuals and in the social system in order to understand individual learning, collaborative knowledge building, and their interplay – regardless of which particular social software tool people may use.

Acknowledgments

The first author of this paper is indebted to the LANDESSTIFTUNG Baden-Württemberg (Königstr. 46, 70173 Stuttgart, Germany) for financial support of this research project by the Eliteprogramme for Postdocs.

References

- Baytiyeh, H., & Pfaffman, J. (2010). Volunteers in Wikipedia: Why the community matters. *Educational Technology & Society*, 13(2), 128-140.
- Bryant, T. (2006). Social software in academia. *Educause Quarterly*, 29(2), 61-64.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*, Cambridge: MIT Press.
- Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers & Education*, 52(1), 141-146.
- Cress, U., & Kimmerle, J. (2007). A theoretical framework for collaborative knowledge building with wikis: A systemic and cognitive perspective. In Chinn, C., Erkens, G., & Puntambekar, S. (Eds.), *Proceedings of the 7th Computer Supported Collaborative Learning Conference* (pp. 153-161), New Brunswick: International Society of the Learning Sciences, Inc.
- Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning*, 3(2), 105-122.
- Cress, U., Kimmerle, J., & Hesse, F. W. (2006). Information exchange with shared databases as a social dilemma: The effect of metaknowledge, bonus systems, and costs. *Communication Research*, 33(5), 370-390.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*, Helsinki: Orienta-Konsultit Oy.
- Engeström, Y. (1990). *Learning, working and imagining: Twelve studies in activity theory*, Helsinki: Orienta-Konsultit Oy.
- Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133-156.
- Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers & Education*, 50(2), 491-498.
- Fessakis, G., Tatsis, K., & Dimitracopoulou, A. (2008). Supporting “learning by design” activities using group blogs. *Educational Technology & Society*, 11(4), 199-212.
- Forte, A., & Bruckman, A. (2006). From Wikipedia to the classroom: Exploring online publication and learning. In *Proceedings of the 7th International Conference on Learning Sciences* (pp. 182-188), Mahwah, NJ: International Society of the Learning Sciences.
- Gibbs, R. W. (2006). *Embodiment and cognitive science*, Cambridge: Cambridge University Press.
- Goldspink, C. (2010). Normative behaviour in Wikipedia. *Information, Communication and Society*, 13(5), 652-673.
- Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5-26.
- Hewitt, J., & Scardamalia, M. (1998). Design principles for distributed knowledge building processes. *Educational Psychology Review*, 10(1), 75-96.
- Hutchins, E. (1995a). *Cognition in the wild*, Cambridge: MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science*, 19(3), 265-288.
- Kim, H. N. (2008). The phenomenon of blogs and theoretical model of blog use in educational contexts. *Computers & Education*, 51(3), 1342-1352.
- Kimmerle, J., Cress, U., & Held, C. (2010). The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. *Knowledge Management Research and Practice*, 8(1), 33-44.
- Kimmerle, J., Cress, U., & Hesse, F. W. (2007). An interactional perspective on group awareness: Alleviating the information-exchange dilemma (for everybody?). *International Journal of Human-Computer Studies*, 65(11), 899-910.
- Kimmerle, J., Moskaliuk, J., Cress, U., & Thiel, A. (2011). A systems theoretical approach to online knowledge building. *AI & Society*, 26, 49-60.
- Kimmerle, J., Moskaliuk, J., Harrer, A., & Cress, U. (2010). Visualizing co-evolution of individual and collective knowledge. *Information, Communication and Society*, 13(8), 1099-1121.
- Köhler, A., & Fuchs-Kittowski, F. (2005). Integration of communities into process-oriented structures. *Journal of Universal Computer Science*, 11(3), 410-425.

- Kolbitsch, J., & Maurer, H. (2006). The transformation of the web: How emerging communities shape the information we consume. *Journal of Universal Computer Science*, 12(2), 187-213.
- Leuf, B., & Cunningham, W. (2001). *The wiki way. Quick collaboration on the web*, Boston: Addison-Wesley.
- Luhmann, N. (1990). *Essays on self-reference*, New York, Oxford: Columbia University Press.
- Luhmann, N. (1995). *Social systems*, Stanford: Stanford University Press.
- Luhmann, N. (2006). System as difference. *Organization*, 13(1), 37-57.
- Majchrzak A., Wagner C. & Yates D. (2006) Corporate wiki users: Results of a survey. In *Proceedings of WikiSym'06 – International Symposium on Wikis 2006* (pp. 99-104), New York: ACM Press.
- Moskaliuk, J., & Kimmerle, J. (2009). Using wikis for organizational learning: Functional and psycho-social principles. *Development and Learning in Organizations*, 23(4), 21-24.
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2009) Wiki-supported learning and knowledge building: Effects of incongruity between knowledge and information. *Journal of Computer Assisted Learning*, 25(6), 549-561.
- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization science*, 5(1), 14-37.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge creating company: How Japanese companies create the dynamics of innovation*, Oxford: Oxford University Press.
- Nonaka, I., & Toyama, R. (2003). The knowledge-creating theory revisited: Knowledge creation as a synthesizing process. *Knowledge Management Research & Practice*, 1(1), 2-10.
- Parker, K. R., & Chao, J. T. (2007). Wiki as a teaching tool. *Interdisciplinary Journal of Knowledge and Learning Objects*, 3, 57-71.
- Piaget, J. (1970). Piaget's theory. In Mussen, P. H. (Ed.), *Carmichael's manual of child psychology* (pp. 703-732), New York: Wiley.
- Piaget, J. (1977). The development of thought: Equilibration of cognitive structures, New York: The Viking Press.
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*, Chicago: University of Chicago Press.
- Polanyi, M. (1966). *The tacit dimension*, Garden City, NY: Doubleday.
- Reinhold, S. (2006). Wikitrails: Augmenting wiki structure for collaborative, interdisciplinary learning. *Proceedings of WikiSym'06 - 2006 International Symposium on Wikis 2006* (pp. 47-57), New York: ACM Press.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In Smith, B. (Ed.), *Liberal education in a knowledge society* (pp. 67-98), Chicago: Open Court.
- Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences*, 3(3), 265-283.
- Scardamalia, M., & Bereiter, C. (1996). Student communities for the advancement of knowledge. *Communications of the ACM*, 39(1), 36-37.
- Scardamalia, M., & Bereiter, C. (1999). Schools as knowledge-building organizations. In Hertzman, C., & Keating, D. (Eds.), *Today's children, tomorrow's society: The developmental health and wealth of nations* (pp. 274-289), New York: Guilford.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In *Encyclopedia of education* (pp. 1370-1373), New York: Macmillan Reference.
- Shih, W.-C., Tseng, S.-S., & Yang, C.-T. (2008). Wiki-based rapid prototyping for teaching-material design in e-Learning grids. *Computers & Education*, 51(3), 1037-1057.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*, Cambridge: MIT Press.
- Stahl, G. (2009). *Studying virtual math teams*, New York: Springer.
- Tsiatsos, T., Andreas, K., & Pomportsis, A. (2010). Evaluation framework for collaborative educational virtual environments. *Educational Technology & Society*, 13(2), 65-77.
- Wang, C., & Turner, D. (2004). Extending the wiki paradigm for use in the classroom. *Proceedings of the International Conference on Information Technology: Coding Computing* (pp. 255-259), Las Vegas: IEEE.
- Yukawa, J. (2006). Co-reflection in online learning: Collaborative critical thinking as narrative. *International Journal of Computer-Supported Collaborative Learning*, 1(2), 203-228.

A Constructivist Approach for Digital Learning: Malaysian Schools Case Study

Waleed H. Sultan, Peter Charles Woods and Ah-Choo Koo

Faculty of Creative Multimedia, Multimedia University, Cyberjaya 63100, Malaysia // waleedsultan@hotmail.com // p.woods@mmu.edu.my // ackoo@mmu.edu.my

ABSTRACT

The purpose of this study was to investigate the influences of Constructivist Learning Environments (CLEs) through the use of laptops supported within 1:1 e-learning education in Malaysian schools. The main objectives of this study were to investigate (a) different possible gaps between constructivist theory and classroom practices in Malaysian schools, (b) success, if any, of the classroom students who would undertake Classmate PC (CMPC) classroom putting while into practice a constructivist approach to learning, and (c) develop a conceptual framework model based on students' communication in a constructivist learning environment. Yet, (d) there is a strong need to address Constructivist Learning Environments (CLEs) practices on local Malaysian settings. A modified Constructivist eLearning Environment Questionnaire (CLEQ) survey used in this study was multiply regressed against student Perceived Learning Outcomes (PRCVD). Findings revealed different aspects of students' learning outcomes and enforcement to use creative thinking in building students' knowledge within constructivism learning context.

Keywords

Constructivist, learning, environments, classroom, process

Introduction

Recent research concerning classroom learning environment has focused on assessment and improvement of learning within the context of Constructivist Learning Environments (CLEs) (Aldridge, Barry, Peter, & Chung-Chih, 2000; Taylor & Cox, 1997; Taylor, Fraser, & White, 1994). A constructivist approach views knowledge as constructed by and embedded in each learner, not something "outside" a learner (Lund & Tannehill, 2010). Learning is not only an "idiosyncratic experience", but it is influenced by social and interaction process. In this study, the researcher attempted to measure the factors of CLEs that might have influenced the "communication between students" and "their learning outcomes". Measure of the factors influencing CLEs in this study was modified from the original version of the study conducted by (Taylor, Fraser, & White, 1994) to suit Malaysian students' age and maturity based on their educational level (age 8-13 years). This study provides an instrument validation for assessing students' preferences toward CLEs and identifies and measures students' perception from five different factors pre-identified from previous research.

Background of the study, ICT initiative in Malaysia school education

The concept of Information & Communication technology (ICT) comes as a part of the Malaysian Information Technology (IT) agenda that exposes all field players starting from students, teachers, administrators and parents to fully utilize IT in every aspect of education at the administrative and classroom levels. School classrooms in this program feature such technology-enablers as laptops personal computers, multimedia computer laboratories, video conferencing systems and high-speed Internet connections. This is done by providing each student with a laptop, providing teacher training to promote project based learning as well as principal training on ICT implementation and development plan.

Related literature

The below sections reviewed related literature to the conceptual framework of the study (Figure 1). Students' perceived e-learning outcomes is mediated by a complex interaction of variables which are: (a) E-Learning resources that needed to link pedagogical curriculum and learning process in the educational environment, (b) a contribution of CLEs dimensions expected to expose students' different activities and engagement in meaningful interactions, (c) possible contribution of different barriers or dilemmas towards integrating CLE such as conceptual, pedagogical, cultural, and political barriers, and (d) students' learning outcome which derived from students' achievements and include classroom activities and performance, academic achievement, and student understanding.

A conceptual framework of this study is based on the integration of different learning theories. Constructivism states that learning takes place in contexts, and that learners form or construct much of what they learn and understand as a function of their experiences in situation (Schunk, 2000). Individuals in an e-learning environment based on constructivist views are forced to use creative thinking to build their knowledge base (Woo, et al., 2007) for meaningful interpretation and reflection of knowledge. (Fosnot, 1996; Schunk, 2008) asserts knowledge is physically constructed by learners who are involved in active learning. This active learning and authentic tasks requires occasionally minimal guidance to enhance student learning and motivation to transfer what is learned to novel problems encountered elsewhere by reducing student cognitive load during tasks (Kirschner, Sweller, & Clark, 2006). Many scholars prefer direct instructions practices in regard to constructivist-based guidance as effectiveness learning strategy (Klahr, 2005; Kolb, et al., 2007).

The constructivist paradigm views the ‘context’ in which learning occurs as central to the activity of learning itself, and this has proved to be a useful theory for designing and developing e-learning programs (McMahon, 2007). Contextualizing learners’ existing knowledge to facilitate the process of new learning experiences is deemed crucial. As Zajda (2004) asserts, cultural and social capital is a crucial dimension in understanding educational processes and academic achievement. In addition, “different social classes maintain different or even antagonistic relations to culture, depending on the conditions in which they acquired their cultural capital and the markets in which they derive most profit from it.” (p. 84). Moreover, Bennett & Desforges (1988) state that a critical factor underlying unsuccessful task implementation is a lack of alignment between a task and students’ prior knowledge, interests, and motivation. It is argued that “the interpretation of knowledge is dependent upon the cultural and social context through which the knowledge is constructed.” (Hung, 2001).

Conceptual framework

Figure (1) shows the conceptual framework for the study that has been developed from literature review. The interactive effects of a large number of variables or concepts and a review of appropriate contextualized methodologies have been also demonstrated. It posits that the students’ perceived e-learning outcomes is mediated by a complex interaction of constructivism learning environment factors, learning environment, 1:1 e-learning resources, learning environment barriers.

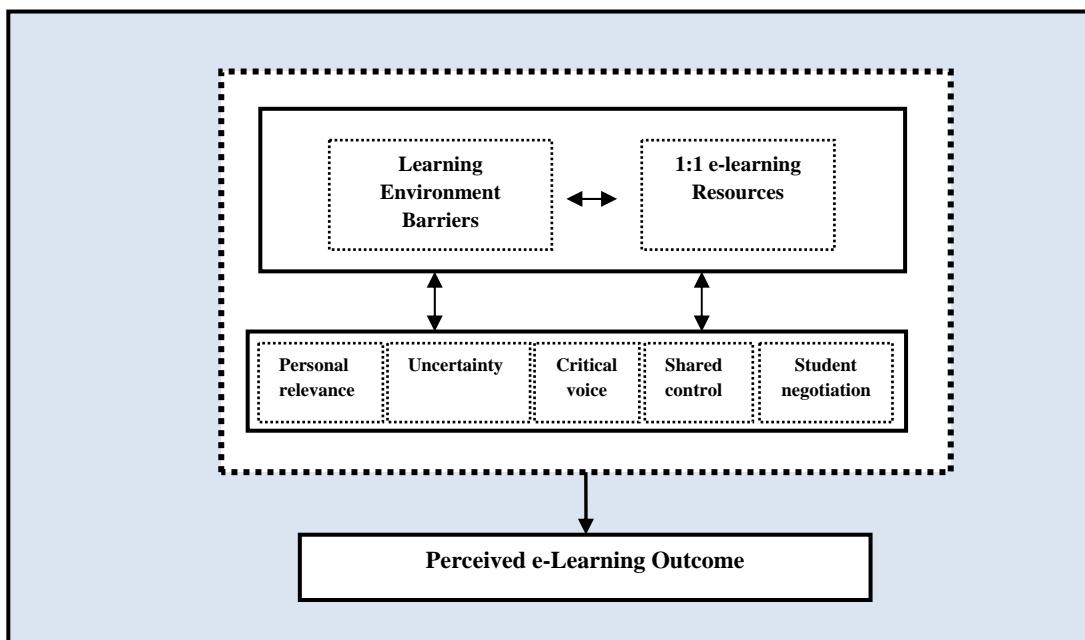


Figure 1: Operational and theoretical definitions of the main research concepts. Adapted from (Taylor, Fraser, & Fisher, 1997; Windschitl, 2002)

E-learning resources

In this study, the term ‘e-learning resources’ refers to learning premises and platforms that include technology enablers such as laptops, multimedia, CD-ROMs, and Internet resources. E-learning resources can give students authentic as well as up-to-date information that is not necessarily available in textbooks. Generally, e-learning resources are useful as they represent a collection of cultural and scientific knowledge accumulated over the years (Yeo, 2008). Moreover, this type of resource can be useful to students because it can foster their learning and their critical thinking, their ability to make connections between different concepts, and bridge the gap between their theoretical and practical knowledge (Palmer, 2007). Despite this, availability of resources does not lead to automatic learning improvement; in fact, productive use of resources can be difficult to achieve (Bera & Liu, 2006).

In Malaysian MOE-INTEL schools, teachers have access to an e-learning system design based on a selected pedagogical model which enables them to make use of the learning resources in a form which is appropriate to the learning goals and the particular learning styles of the students (Granic & Cukusic, 2007). Constructivist theory emphasizes the need for student-centred learning rather than teacher-centred learning (Kelsey, 2007; Young & Maxwell, 2007). Constructivist e-learning emphasizes that there are several learning activities and development programmes for students. This could be clearly in the deployment of online resources such as online information and communication resources that are useful for effective learning activity (Paurelle, 2003). The practical implication of making online resources available to students is to link students’ knowledge to other useful online resources and facilitating the learners’ journey of discovery and acquisition of new knowledge. Communication resources such as discussion boards enable learners to participate in collaborative learning with other students and with educators. Students can share ideas at anytime from anywhere through the online course.

Constructivist Learning Environments Dimensions

CLEs are technology-based in which students are expected to expose different activities and engagement in meaningful interactions (Jonasson, 1999). Students in CLEs should be provided with opportunities to negotiate ideas, conduct inquiry, and reflect their thoughts towards enhancing their cognitive and metacognitive outcomes. Taylor and his colleagues (1994, 1996) had conducted several CLEs studies a related to the measure of CLEs. It is believed that, the ICT tools such as 1:1 learning resources and its institutional approaches can bring and concretize the ideas of the Constructivist Learning Environments as mentioned by different scholars (Taylor, Fraser, 1991; Johnson & McClure, 2000; McClure & Gatlin, 2007). These factors are described as follows:

1. Personal relevancy of the students is important in the sense that it concerns with the connectedness of school [experiences] to students’ out-of-school experiences, and on making use of students' everyday experiences as a meaningful context for the development of students' knowledge. As (Selvi 2007) asserted “*Science curriculum must be developed based on student and society needs, scientific and technological developments in the field of science and educational science..*”.
2. Student uncertainty involves the extent to which opportunities are provided for students to experience scientific knowledge as arising from theory-dependent inquiry involving human experience and values, and as evolving, non-foundational, and culturally and socially determined. Abma & Widdershoven, (2006) stated “*Central feature of dialogue are openness, respect, inclusion and engagement.*”, while a great enhancement in scientific literacy is the explicit goal of science educators (von Aufschnaiter, Erduran et al. 2008).
3. Critical voice of the students involves the extent to which a social climate has been established in which students feel that it is legitimate and beneficial to question the teacher's pedagogical plans and methods, and to express concerns about any impediments to their learning and find room and other ways of knowing. This is clearly argued as Moore, (2006) asserted “*students' voices when critically reflect how they viewed their instructional practices can help to affirm students' to engage and contribute to positive changes in the classroom.*”
4. Students’ shared control concerns students being invited to share with their teacher controlling of the learning environments in a sense of articulation of learning goals, design and management of learning practices, and the determination and application of assessment criteria. Earlier, Vygotsky (1978) stated “*through practical activity a child constructs meaning on an intrapersonal level, while speech connects this meaning with the interpersonal world shared by the child and her/his culture*”. The greater perceived control offered by shared control will

- reveal positive effects on students' motivation and reveals higher learning outcomes leading to higher performance (Moores & Chang, 2009).
5. The concept of students' negotiation is one of the most influential perspectives that directly affect students' knowledge, ideas, attitudes, and values when these students interact to each other's (Vygotsky, 1978). Students' negotiation considered as a key element in building classroom learning environment (Davis, 2006). A constructivist approach to learning is based on the idea that the learner constructs his or her own knowledge through negotiation of meaning (Hand, Treagust et al. 1997), which will confirms the need for cognitive performance of students in classroom is related to their interaction (Taylor and Cox 1997; Barron, Kemker et al. 2003). (Smith, Maclin et al. 2000).

Barriers of integrating constructivist learning environment in classroom

The school is considered to be a component of society. Yet, possible emerging pedagogies could be integrated in the public domain through schools. Schools now require teachers to place great concerns that related to communication skills, paths of students' personal learning, and student self-evaluation. Constructivism in practice may be defined as multi-interactions of different activities and contexts of teaching that bind together students, teachers, administrators and community as all of them participate in pedagogy.

Within schools' settings, different aspects might work against practicing applications of the theoretical ideals of constructivism. These barriers could include: (1) conceptual barriers which arise when there is a need to acquire new dimensions of instructional expertise that "are rooted in teachers' attempts to understand the philosophical, psychological, and epistemological underpinnings of constructivism;" (2) pedagogical barriers which arise from the need for teachers to develop more complex approaches to designing instructional materials that constructivist learning and the constructivist curriculum require because this necessitates that teachers work hard, concentrate more and embrace larger pedagogical responsibilities (Cohen & Educational Resources Information, 1988); (3) cultural barriers which emerge in the constructivist classroom involving teachers and students and which require an understanding of the norms and values necessary to accommodate the constructivist approach as it's crucial to understand and consider multidimensional cultural realities in school before implementing curriculum and pedagogical proposals (Zajda, 2004); and (4) political barriers which are associated with resistance from various environments outside the school and which require teachers to deal with. In the same vein, the constructivist 1:1 learning practice endorse a complicated concerns binds policy makers, teachers, students, and parents of students as they all participate in reforming the concept of teaching in general, and teaching for understanding in particular, in the school setting.

Learning Outcomes

Learning outcomes is the measurable cognitive dimension that occurs through the learning process. Students' cognitive processes refer to "students learning skills, critical thinking, high-order cognitive process, problem-solving capability, knowledge and skills in the key learning areas and the ability to transfer knowledge or skill." Also, "the affective domain includes attitudes, motivation and goal structures, and self-perceptions of ability or skill that include self-concept, self-esteem and self-efficacy". Initial learning goals are set by the instructor, but students have the chance to reflect learning that matches their individual interest (Driscoll, 2005; Johnson, 2000). Similarly, different measures could be used to measure affective student's outcomes in a constructivist learning as: "self-report measures", "classroom observations" and "analysis of student performance data". Some researchers (e.g. Hase, 2000; Woo & Kimmick, 2000) argue that the quality of learning materials is an important factor influencing the learning outcomes of students. It should be pointed out, however, that measuring students' learning outcomes is beyond the scope of this study.

Research hypotheses

The hypotheses for this study are based on CLEs conceptual framework and the measures as described above:

H1: Students' 'Personal relevance' understanding of constructivist practices using e-learning resources is positively related to their 'Perceived e-learning outcome'.

H2: Students' 'Uncertainty' understanding of constructivist practices using e-learning resources is positively related to their 'Perceived e-learning outcome'.

H3: Students' 'Critical voice' understanding of constructivist practices using e-learning resources is positively related to their 'Perceived e-learning outcome'.

H4: Students' 'Shared control' understanding of constructivist practices using e-learning resources is positively related to their 'Perceived e-learning outcome'.

H5: 'Student' negotiation' understanding of constructivist practices using e-learning resources is positively related to their 'Perceived e-learning outcome'.

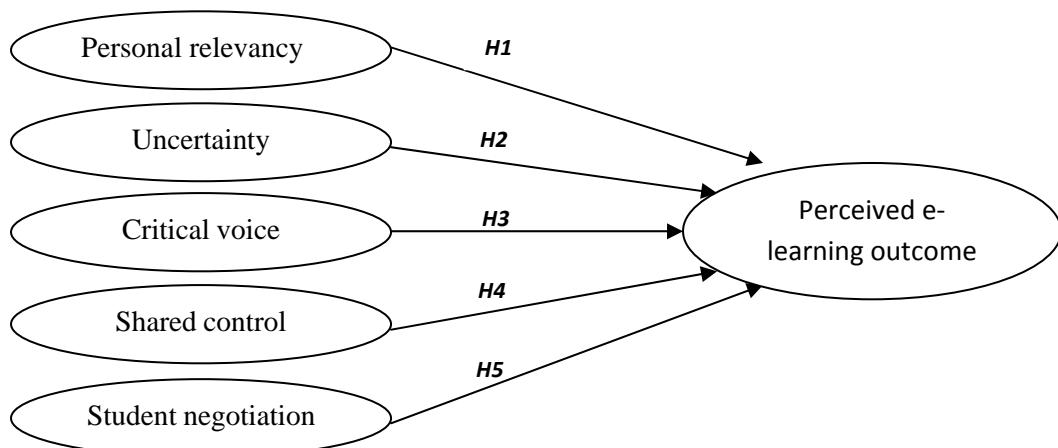


Figure 2: Hypothetical CLEs Model

Research Method

Overview

The study used quantitative research methods. The purpose of quantitative research is to describe phenomena numerically and to answer specific research questions. The aim of this study was to assess students' perceptions of their classroom experiences with 1:1 learning of computing and their understanding of the dimensions of Constructivist Learning Environments (CLEs) identified in this study. The study elicited data from Malaysian students from MOE-Intel schools which were utilizing in 1:1 learning resources through the introduction of Intel classmate PC (CMPC) program. A new Malaysia-version of the Constructivist Learning Environment Questionnaire (CLEQ) was developed to provide researchers or teachers with important information that could help improve the quality of the teaching and learning process. The internal consistency of CLEQ was examined using Cronbach's Alpha values (Table 1). Constructivist Learning environment (CLE) factors ("Personal relevance", "Uncertainty", "Critical voice", "Shared control", and "Students' Negotiation") was multiply regressed against student perceived learning outcomes (PRCVD) to identify level of classroom practices in a constructivist learning environment.

Methods and sample

The population of this study was Malaysian students who utilizing in 1:1 learning resources through the introduction of Intel classmate PC (CMPC) program. The questionnaire survey contains demographic questions and different parts of evaluations segmented as: (a) an evaluation of students' personal relevance of e-learning, (b) their uncertainty of science subjects, (c) their experience to speak out, (d) their shared control, (e) their learning negotiation, and (f) their assessments of perceived e-learning outcomes. In general, seven high schools and three primary schools were chosen by Malaysian MOE as a 1st phase to apply 1:1 e-learning Project (MOE report, 2007). Target students for this study were computer literate; able to speak in English; had exposure to different e-

learning resources; and able to access Internet. In addition, classroom is equipped with Wi-Fi connection; students' seats were allocated to facilitate groups' discussion, and equipped with small laptop for each student named Classmate PC (CMPC). From 10 selected schools of 10 different states, a total of 304 out of 608 students were randomly selected to participate in 'Constructivism Learning Environment Questionnaire (CLEQ) survey. After cleaning the data, this resulted in 291 usable questionnaires. Thus, 96% usable responses rate was achieved.

Analysis of the measurement of the constructs

The construction of CLEs Model was accomplished among several stages: (a) exploratory list of terms identified as: relevancy, uncertainty, critical voice, shared control, and student negotiation, (b) content validity which had been employed by careful selection of dimensions based on related literature revision and suggested methods for constructs and scale development obtained from pretested studies (Churchill, 1978, 1999), (c) construct validity with implementing discriminant validity that showed correlation among constructs are less than 0.9. (Hair et al., 2006), and convergent validity that, all items had loadings higher than 0.40. Further, convergent validity was ensured with Composite Reliability (CR) of 0.92 and an average variance extracted (AVE) of 0.53 (Hair et al., 2006). (d) Consistency reliability has been evaluated by computing the Cronbach's alpha coefficients for each scale (Table 1). All six of the subscales possess internal reliability of alpha greater than .70. The Kaiser-Meyer-Olkin (KMO) value is .805, and Bartlett's Test of Sphericity indicates that the significance level (sig) for Sphericity (2263.059) for the six-item correlation matrix is highly significant ($p < .001$). The significance value is $p < .05$, which confirms that this data is suitable for Factor Analysis. (Table 1). Factor loadings of each measurement-item and their corresponding components were generated (Table 1).

Table 1: KMO & Bartlett's Test

| | |
|--|----------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .805 |
| Bartlett's Test of Sphericity | |
| Approx. Chi-Square | 2263.059 |
| df | 276 |
| Sig. | .000 |

From Table: 2, all factors-items has a loadings and a cut-off point of 0.40 and no significant cross loadings criteria were identified (Hair et al., 2006; Tabachnick & Fidell, 2000). This will further suggest that, CLEs confirms six factors corresponding to (RELV, UNCE, CRTVC, SHRCNT, NEGOT, and PRCVD).

Table 2: CLEs list of terms, Loadings, CA and AVE (Adapted from Taylor, Fraser, 1991; Johnson & McClure, 2000; McClure & Gatin, 2007)

| Learning about the world (relevancy) | | | Loadings |
|--|-------|--|----------|
| CA = 0.740 AVE = 0.566 | A.1.1 | 1. The things I learn are relevant to me. | .840 |
| | A.1.2 | 2. I am enjoying of new ways of learning when I am studying. | .823 |
| | A.1.3 | 3. The things that I learn about are relevant to my current and/or future educational environment(s). | .663 |
| | A.1.4 | 4. I learn things that are a part of my-out-school life. | .455 |
| Learning about research process (uncertainty) | | | |
| CA = 0.731 AVE = 0.555 | B.1.1 | 5. I learn that classroom activities cannot always provide clear answers to problems. | .788 |
| | B.1.2 | 6. I learn that modern science is different from the science of long ago. | .788 |
| | B.1.3 | 7. I learn that classroom activities can be influenced by people's experiences, opinions, and values. | .698 |
| | B.1.4 | 8. I learn that classroom activities are a way to seek better answers and generate new questions. | .688 |
| Learning to speak out (critical voice) | | | |
| CA = 0.764 AVE = 0.585 | C.1.1 | 9. I feel safe questioning what or how I am being taught. | .782 |
| | C.1.2 | 10. I feel I learn better when I am allowed to question what or how I am being taught. | .777 |
| | C.1.3 | 11. It's OK for me to ask my teacher for clarification about activities that are confusing. | .749 |
| | C.1.4 | 12. My teacher encourages me to ask questions to clarify ideas, or to deepen my | .704 |

| | | | |
|--|-------|---|------|
| | | understanding. | |
| Learning to learn (shared control) | | | |
| CA = 0.765 | D.1.1 | 13. I help to plan what I am going to learn. | .858 |
| | D.1.2 | 14. I help to decide how well I am learning. | .835 |
| | D.1.3 | 15. I help to decide which activities work best for me. | .755 |
| AVE = 0.800 | D.1.4 | 16. I let the teacher know when I need more or less time to complete assignments. | .533 |
| Learning to Share (Student negotiation) | | | |
| CA = 0.750 | E.1.1 | 17. I talk with other students about how to solve problems. | .878 |
| | E.1.2 | 18. I ask other students to explain their ideas. | .781 |
| | E.1.3 | 19. I am asked by other students to explain my ideas. | .687 |
| AVE = 0.573 | E.1.4 | 20. My teacher encourages me to raise issues and ask questions with other students in order to clarify and inform our thinking. | .569 |
| Perceived E-Learning Outcome (Student Experience) | | | |
| CA = 0.767 | F.1.1 | 21. Using the e-learning resources improved my critical thinking | .853 |
| | F.1.2 | 22. The e-learning resources help me understand the subject contents more in-depth. | .820 |
| | F.1.3 | 23. I enjoy using different learning modes to learn. | .810 |
| AVE = 0.60 | F.1.4 | 24. The e-learning resources help me understand the subject contents quicker. | .513 |
| Note: CA = Cronbach Alpha, AVE = Average Variance Extracted | | | |

Results

Table 2 shows the items of each construct (relevancy, uncertainty, critical voice, shared control, students' negotiation and learning outcomes). The descriptive summary and correlation coefficient are provided. The highest value for each construct's item with at least one other item in the construct within 0.3 to 0.9. The factor loadings are provided in the last column of table 2. Each factor score was saved under a given name (RELV, UNCER, CRTVC, SHRCONT, NEGOT, and PRCVD) to be used for further analysis.

Findings

Multiple regressions was used to evaluate the relationship of the constructs with perceived e-learning outcome. Multiple regression aims to use the independent (predictor) variables whose values are known to predict one dependent (criterion) variable (Hair, 2010). Using Stepwise Method, four items were entered (RELV, UNCER, SHRCONT and NEGOT) and one item (CRTVC) failed to meet the criteria (entered). The R² value was 26.5% and adjusted R² was 25.5%. The R² indicates that 26.5% of the variance in P can be explained by REV, UNCER, SHRCONT and NEGOT. In the ANOVA table, the F-Value was 25.8, and the degree of freedom (df) was 4 and residual was 286. The p-Value of the test was less than 0.001. Therefore, P depends on all factors except CRTVC.

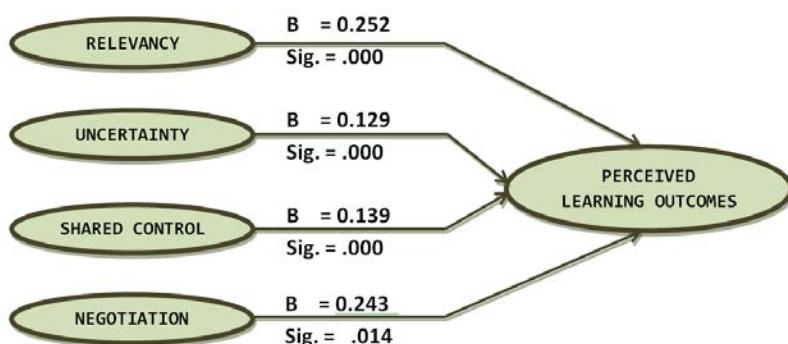


Figure 3: Relationships of CLEs factors with Perceived e-learning outcome

Discussion

This study suggests that, students' CLEs for classroom practice is positively related to their "perceived e-learning outcomes". The "perceived learning outcomes," is a result of student's real interaction of real world problems as Jonassen & Jonassen (2006) believes learners should be presented with interesting, relevant, and meaningful problems to solve. Student's learning experience concerns the level and source of motivation for learning which revealed a workable confidence level to students in his or her potential of learning (Glaserfeld, 1989 #360). This viewpoint chime with CLEs concept that students' constructivist classroom outcomes is a function of their beliefs and practices of e-learning resources and supported with real world problems and improvement in student attitudes toward science (Yore, et al., 2005). The results suggest that "perceived e-learning outcomes" of the students is a function of different factors (Personal relevance, Uncertainty, Critical voice, Shared control, and Students' negotiation). These factors are discussed below.

Personal relevance

The hypothesis result extracted from regression analysis, suggests that students' 'personal relevance' (RELV) has a positive "t-value" (4.067) and a significant p-value ($p < .05$) influence on students' 'perceived e-learning outcome'. About 23.4% of the variance in the 'perceived e-learning outcomes' score can be accounted for by the 'personal relevance' (RELV) score. The findings support the first hypothesis (H1). The study suggests that, students' courses need to have content which has relevance to out of school life. Moreover, students need to practice and experience personal knowledge through reflection, critical thinking, and self-regulated learning, otherwise their learning is only superficial and their way of thinking is not modified (Piaget, 1989). Similarly, Piaget (1989) argues that, "What has not been acquired through experience and personal reflection can only be superficially assimilated and does not modify any way of thinking" (p. 252).

The theoretical and empirical implications of the findings are that, the students' learning outcome could be enhanced in class if students perceive more personal relevance in their studies and are able to express different issues related to their learning and can view science subjects as changing phenomena.

There is still a need for further research on students' personal relevance in relation to knowledge acquisition and learning within Malaysian educational context. The MOE is thus "concerned with different knowledge transfer strategies such as codification and personalization" (Werr & Stjernberg, 2003). However, there is a few debate concerning students' sense-making of professional interacts with the socio-cultural context during their process of in-school and out-of-school knowledge transfer. It is quite important to see the school as an institution that has valuable "effects on knowledge development" (Robertson, Scarbrough, & Swan, 2003), and to define what role the MOE can play in becoming a source of innovation (Nikolova, 2007) that could facilitate students' ability to perceive more personal relevance in relation to science subjects and to understand science phenomena. Yet, the study recommends that further research is needed to examine the extent to which the 'personal relevancy' scale generates a plausible account of "classroom learning environments" that offer extensive opportunities for students to develop their understanding of the relevancy of their in-school learning to that of real-world problems and related activities outside school."

Student uncertainty

The hypothesis result extracted from regression analysis (Figure 3) suggests that students' 'uncertainty' (UNCER) has a positive "t-value" (1.983) and a significant p-value ($p < .05$) influence on their 'perceived e-learning outcomes'. About 10.3% of the variance in the 'perceived e-learning outcomes' score can be accounted for by the students' 'uncertainty' (UNCER) score. The findings support the second hypothesis (H2).

The practical implications of the findings are that students' uncertainty in relation to the constructivist e-learning environment influences their cognitive processes and engagement with activities carried out by scientists (Brill & Yarden, 2003; Chinn & Malhotra, 2002; Lin & Tallman, 2006). Although, scientific inquiry itself is a complex form of human thought and can be considered in a Vygotskian sense as a "cultural tool" (Wertsch, 1991), which Malaysian students should implement as an approach to reasoning that others have found useful. From the socio-

cognitive perspective, three principles underpin these scientific inquiry instructions: (a) engaging students in reasoning practices in science subjects; (b) offering explicit guidance on the roles students can assume to monitor their own and their peers' thinking; and (c) fostering a sophisticated epistemology of science by having students experience science as a process of revision. Within the MOE-INTEL classroom, students are able to share aspects of learning science with their teachers and express their doubts about specific science topics. However, "knowledge only has a lasting effect when it is integrated into a student's cognitive structure in a way that it is properly understood, and appears relevant, useful and important to the student's learning context when built up by assigned members and integrated into their cognitive structure in understandable way" (Reihlen & Apel, 2007).

This study is thus argues that, in order to determine the extent to which items in the 'uncertainty' scale can elicit dependable and clear responses, it is necessary to investigate 'uncertainty' in a classroom learning environment that provides rich and sustained experiences in a predefined science topic whose learning goal has been affirmed by the constructivist teacher. Teachers themselves also need to have an adequate pedagogical and conceptual understanding of constructivist learning."

Critical Voice

The hypothesis result extracted from regression analysis (Figure 3) suggests that students' 'critical voice' (CRTVC) has a positive "t-value" (1.217) but not a significant p-value ($p > .05$) influence on their 'perceived e-learning outcomes'. The findings would appear not to support the third hypothesis (H3). In Malaysian schools, students are expected to be exposed to a different range of media and it is the responsibility of teachers to equip students to engage and learn critically in their classrooms. The practical and theoretical implications of these findings regarding the students' critical voice are that it directly influences their socio-cognitive process. The evidence from the analysis suggests that, in order to enhance the attribute of the sound voice in Malaysian students which is required for their learning environment, the revision of their learning strategy should be enlarged to consider: (1) defining social and cultural presence to facilitate students' ability to feel socio-culturally connected to their classroom and learning environment (Aragon, 2003); (2) providing opportunities for students to interact and share different ideas regarding course content as an effective way to engage students in learning activities; and (3) improving motivation and encouraging bonding between students (Pratt & Palloff, 2007).

Yet, students' "Critical Voice" should be treated as a critical "CLEs" factor that offers extensive opportunities for students to critically judge their learning environments. Significantly, teachers' lack of constructivist epistemological skills may contribute to a major impediment that can distract from student's role in the 'critical voice' dimension of the constructivist learning environment scale. Further research is thus needed on the critical voice of students, the focus of which should be broadened in the context of the enculturation of constructivist pedagogy in Malaysian schools.

Shared Control

The hypothesis result extracted from regression analysis (Figure 3) suggests that students' 'shared control' (SHRCONT) has a positive "t-value" (2.581) and a significant p-value ($p < .05$) influence on their 'perceived e-learning outcomes'. About 13.9% of the variance in the 'perceived e-learning outcomes' score can be accounted for by the 'shared control' (SHRCONT) score. The findings support the fourth hypothesis (H4). There is a need to effectively develop shared control strategies in the learning environment that will allow students to be adept members of today's multi-literate society (McNaughton, 2002). However, the students' role in these aspects has too often been reduced. Students should be educated to utilize their capability to question their teachers' on their pedagogical plans. Additionally, students within a "CLEs" should be invited and encouraged to share with their teachers' articulation of learning goals; designing and managing their learning activities; and be able to determine different assessment criteria for their learning activities. The practical implication of the findings of this study is that, acting of the constructivism learning teacher as a facilitator will provide students with variety of learning experience from which learning is built and this will foster social interactions between students themselves towards negotiate meaning of their learning (Brooks & Brooks, 1999). Similar findings also had been argued Nunes & McPherson (2003) who asserts that "Students are expected to develop high cognitive skills such as negotiation of meaning and metacognition supported by teachers who often may lack these skills themselves" (P.

146). Also, role of constructivist classroom teacher should concern facilitating the process of teaching-and-learning in away that locus of the control for learning is passed from the teacher to the learner in a careful guided way. As students' "self-efficacy" will increase (Hase, 2000), and the amount of effort and time teacher and student is willing to devote to the task also increases, leading to higher performance (Moores & Chang, 2009). Generally, the greater perceived control offered by shared control, the greater the positive effects on students' motivation, demonstrated by increased task involvement, which in return will lead to higher learning outcome.

Students' Negotiation

The hypothesis result extracted from regression analysis (Figure 3) suggests that student negotiation (NEGOT) has a positive "t-value" (4.714) and a significant p-value ($p < .05$) influence on their 'perceived e-learning outcomes'. About 26.9% of the variance in the 'perceived e-learning outcomes' score can be accounted for by the 'student negotiation' (NEGOT) score. The findings support the fifth hypothesis (H5). The concept of student negotiation is one of the most influential perspectives that directly affects students' knowledge, ideas, attitudes, and values when these students interact with each other (Vygotsky, 1978). The affective interaction of students in their learning environment places the students' relationship with each other in the foreground because negotiation, cooperation, conflict, rhetoric, roles, etc. are important as means by which students are able to construct knowledge (Gergen, 2001). The implication of student negotiation of meaning in the domain of education confirms the need for cognitive frames of references as the performance of students in the classroom is related to team interaction (Barron, 2003). Similarly, students will have greater opportunities to negotiate when they have a trusting relationship with their classroom teacher, which in fact is considered a key element in building a classroom learning environment (Davis, 2006; Davis & Cosenza, 1993; Davis, 2006; Turner et al., 1998; Turner et al., 2002).

Perceived e-learning outcome (student outcome)

The results from the quantitative data collected suggest that four hypotheses are supported, which are related to: personal relevancy; uncertainty; shared control; and student negotiation. One hypothesis is not related: (H3) students' 'critical voice' is positively related to their 'perceived e-learning outcomes' ($p > 0.05$). To sum up, the findings of this study revealed that students who perceived their learning environment in a more constructivist-oriented manner believed that knowledge is changing and that they have ample opportunity to think critically and adopt meaningful learning strategies.

The students' 'perceived e-learning outcome' is positively related to four dimensions of CLEs. The results derived from the conceptual framework of the study (Figure 2), and the multiple regressions analysis, represented in the Constructivist Learning Environment Model (CLEM), (Figure 3), and hypothetical results, suggest four hypotheses are supported, which are: (H1) students' 'personal relevance' understanding of constructivist practices using e-learning resources significantly influences 'perceived e-learning outcomes' (t-value 4.067, p-value $p < .05$); (H2) students' 'uncertainty' understanding of constructivist practices using e-learning resources significantly influences 'perceived e-learning outcomes' (t-value 1.983, p-value $p < .05$); (H4) students' 'shared control' understanding of constructivist practices using e-learning resources significantly influences 'perceived e-learning outcomes' (t-value 2.581, p-value $p < .05$); and (H5) students' 'student negotiation' understanding of constructivist practices using e-learning resources significantly influences 'perceived e-learning outcomes' (t-value 4.714, p-value $p < .05$).

Some researchers (e.g. Hesse, 2004; Woo & Kimmick, 2000) argue that, the quality of learning materials is an important factor influencing the learning outcomes of students. A learning environment where students are able to contribute to their acquisition of their knowledge through experimentation, observation, raising their voices, questioning, and negotiating with their teachers and peers can help students realize that scientific knowledge is evolving, and they are then apt to enter an enriched new learning environment (Hanson-Smith, 2000). Similarly, Cano (2005) highlights the importance of the relationship between the learning environment and students' metacognition and epistemological beliefs. He asserts that,

"A student's academic performance or achievement will depend on his/her approach, which in turn will depend on other factors. These are characteristics of the teaching context (teaching styles, ways of assessment, etc.) and the pupil's perception about this context (learning environment) and the

academic tasks set (metacognition), which is also related to his/her personal characteristics (intellectual skills, personality, etc.), including beliefs about knowledge and learning (epistemological beliefs)" (p. 205).

Overall, students' understanding of the five dimensions of CLEs establishes a positive relationship between the students' constructivist learning environment and their perceived e-learning outcomes.

Based on the constructivist assumption that students need to feel the ownership of their knowledge construction, future research is needed to evaluate students' beliefs regarding their contribution to their learning outcomes. Students' perceived learning outcome could be a subject of further study that could investigate the extent to which students' beliefs in their learning environment, and their intention to fully engage in understanding and utilization of "CLEs" dimensions as well as their reaction to values, affect the quality of their self-regulatory learning

Future Research

Future research is needed on one constructivism learning strategy (e.g. Problem based learning) and how they use it in-and-out-of Malaysian schools. However, integration of ICT in school's pedagogy and students' computer literacy towards their learning outcome in a constructivism learning environment could be another dimension to be considered. Moreover, intervention studies can be conducted to determine connections between teacher's pedagogical practices and conservative beliefs and attitudes in a constructivist learning practices. Finally, Gender differences in classroom practices and learning outcomes is also possible when learners are required to construct their own learning. To which extend does the gender will vary when it comes to this concept? As Hoffman (2008) argued "*While the trends indicate both sexes are using social media in huge numbers, our findings show that women far outpace the men*". Does this could be possible in Malaysian schools when applying 1:1 learning environment strategy?

Conclusion

The dimensions of Constructivist Learning Environments are the practical measurable constructs that were used in this study. A review of various studies shows that the majority of early studies on the learning environment were conducted in Western countries; however, over the last decade, Asia-based researchers have made important contributions to this field (Fraser, 1998). Studies conducted in Indonesia (Margianti, 2001), Singapore (Chionh & Fraser, 2009), Korea (Kim, Fisher, & Fraser, 1999), and Brunei (Scott, 2001) have replicated prior research findings, namely that the psychosocial aspects of the learning environment are an important determinant of student outcomes. Despite the fact that a great deal of learning environment research has been conducted all over the world, there are few related works that have been undertaken in regard to the constructivist learning environment in Malaysian schools. Different studies (Taylor, & Fraser, 1991; Fraser et al., 1997) have conducted investigations into several CLEs and measures as well as on the introduction and use of 1:1 learning resources in the Malaysian MOE-Intel classroom.

The main contribution of this research is that it adds to our knowledge and understanding of the concept of constructivism in education. The contributions of the study are: (1) the concept of constructivist learning is grounded with accessible pedagogies that frame constructivist learning. This study identified and compiled literature that focuses on CLEs within the Malaysian school context that emphasizes the integration of multimedia technology in a 1:1 learning environment, and identified different barriers teachers may face when attempting to implement the concept, (2) the theoretical framework of this study identified different barriers impeding the implementation of constructivist learning within Malaysian schools, and how it prevent implementation of constructivist learning that promotes students' development and encourages positive learning outcomes, (3) this study identified a new instrument CLEQ and proposed as a Malaysia-version, providing teachers or researchers with information that could assist in improving the ultimate quality pertaining teaching and learning process within a constructivist learning, (3) based on both theory and empirical findings, the study proposed a Constructivist Learning Environments Model (CLEM) that explains how using different constructivist learning factors "Personal Relevance", "Uncertainty", "Critical Voice", "Shared Control", and "students' Negotiation" in 1:1 e-learning influences students' perceived e-learning outcomes, and (4) by taking into account the Malaysian MOE school

reform strategy, namely, the MOE-Intel project, this study also addressed the context of constructivist learning within a 1:1 learning environment that is characterized by the integration of ICT and multimedia technology into education and laptops for students (CMPC). The findings from this study contribute to the growing body of knowledge on the learning experience in Malaysian schools and specifically in respect of the 1:1 learning environment through the example of the MOE-Intel project.

References

- Abma, T., & Widdershoven, G. (2006). Moral Deliberation in Psychiatric Nursing Practice. *Nursing Ethics*, 13(5), 546-557.
- Aldridge Jill, M., Barry, J. F., Peter, C. T., & Chung-Chih, C. (2000). Constructivist learning environments in a crossnational study in Taiwan and Australia. *International Journal of Science Education*, 22(1), 37.
- Aragon, S. (2003). Creating Social Presence in Online Environments. *New Directions for Adult and Continuing Education*, 2003(100), 57-68.
- Barron, A. E., Kemker, K., Harmes, C., & Kalaydjian, K. (2003). Large-Scale Research Study on Technology in K-12 Schools: Technology Integration as It Relates to the National Technology Standards. *Journal of Research on Technology in Education*, 35, 489-507.
- Barron, B. (2003). When Smart Groups Fail. *The Journal of the Learning Sciences*, 12(3), 307-359.
- Bera, S., & Liu, M. (2006). Cognitive tools, individual differences, and group processing as mediating factors in a hypermedia environment. *Computers in Human Behavior*, 22(2), 295-319.
- Brill, G., & Yarden, A. (2003). *Learning Biology through Research Papers: A Stimulus for Question-Asking by High-School Students*. Retrieved May 1, 2011 from <http://www.lifescied.org/cgi/reprint/2/4/266.pdf>.
- Brooks, J. G., Brooks, M. G., & Association for Supervision and Curriculum Development (1999). *In Search of Understanding. The Case for Constructivist Classrooms [Revised.]*, Alexandria, VA: Association for Supervision and Curriculum Development.
- Cano, F. (2005). Epistemological beliefs and approaches to learning: Their change through secondary school and their influence on academic performance. *British Journal of Educational Psychology*, 75(2), 203-221.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education*, 86, 175-218.
- Chionh, Y. H., & Fraser, B. J. (2009). Classroom Environment, Achievement, Attitudes and Self-Esteem in Geography and Mathematics in Singapore. *International Research in Geographical and Environmental Education*, 18(1), 29-44.
- Churchill, G. A. (1978). *A paradigm for developing better measures of marketing constructs*. [Madison]: Graduate School of Business, University of Wisconsin-Madison.
- Churchill, G. A. (1999). *Marketing research : methodological foundations*. Fort Worth: Dryden Press.
- Cohen, D. K. (1988). *Teaching practice plus ca change*. East Lansing, Mich.: Michigan State University.
- Davis, B. M. (2006). *How to teach students who don't look like you : culturally relevant teaching strategies*. Thousand Oaks, CA: Corwin Press.
- Davis, D., & Cosenza, R. M. (1993). *Business research for decision making*. Belmont, CA: Wadsworth.
- Davis, H. A. (2006). Exploring the Contexts of Relationship Quality between Middle School Students and Teachers. *The Elementary School Journal*, 106(3), 193.
- Driscoll, M. P. (2005). *Psychology of learning for instruction*. Boston: Pearson Allyn and Bacon.
- Florio, S. E. (1978). *Learning how to go to school : an ethnography of interaction in a kindergarten/first grade classroom*, Harvard: Harvard Graduate School of Education.
- Fosnot, C. T. (1996). *Constructivism: theory, perspectives, and practice*. New York: Teachers College Press.
- Fraser, B. J. (1998). Classroom Environment Instruments: Development, Validity and Applications. *Learning Environments Research*, 1(1), 7.
- Gergen, K. J. (2001). Psychological science in a postmodern context. *The American psychologist*, 56(10), 803-813.
- Granic, A., & Cukusic, M. (2007). Universal Design Within the Context of e-Learning. *Lecture notes in computer science*, 4556, 617-626.

- Hair, J. F. (2006). *Research methods for business*. Chichester: John Wiley.
- Hair, J. F. (2010). *Multivariate data analysis : a global perspective*. Upper Saddle River, N.J.; London: Pearson Education.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with readings* (Vol. 103-388-996): Prentice Hall (Englewood Cliffs, NJ)
- Hand, B., Treagust, D. F., & Vance, K. (1997). Student Perceptions of the Social Constructivist Classroom. *Science Education*, 81(5), 561.
- Hanson-Smith, E. (2000). *Technology-enhanced learning environments*. Alexandria, VA: Teachers of English to Speakers of Other Languages.
- Hase, S. (2000). Capability and coping in the new millennium: a new challenge for education. *Paper presented to Universities, Colleges and Sustainable Health: a 21st Century Investment Conference*, 4-6 September, Preston, UK.
- Hase, S. K., C. (2000). From andragogy to heutagogy. Retrieved February, 7, 2010, from <http://ultibase.rmit.edu.au/Articles/dec00/hase1.pdf>.
- Hesse, G. (2004). *Proceedings of the International Conference of the Learning Sciences*. Mahwah: Lawrence Erlbaum.
- Hung, D. (2001). Theories of Learning and Computer-Mediated Instructional Technologies. *Educational Media International*, 38(4), 281-287.
- Iliff, C. H. (1994). *Kolb Learning Style Inventory: a meta-analysis*. Unpublished Doctoral dissertation, Boston University.
- Johnson, B., & McClure, R. (2000). How are our graduates teaching? Looking at the learning environments of our graduates' classrooms. *A paper presented at the annual meeting of the Association for the Education of Teachers in Science*.
- Jonassen, D. H., & Jonassen, D. H. (2006). *Modeling with technology : mindtools for conceptual change*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Kelsey, K. (2007). *Constructivism: strength and weakness*. Retrieved July 13, 2011, from <http://adulteducation.wikibook.us/index.php?title=Constructivism>.
- Kim, H.-B., Fisher, D. L., & Fraser, B. J. (1999). Assessment and Investigation of Constructivist Science Learning Environments in Korea. *Research in Science & Technological Education*, 17(2), 239.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance during Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), 75-86.
- Klahr, D., & Nigam, M. (2004). *The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning*. Retrieved May, 9, 2011, from <http://www.psy.cmu.edu/faculty/klahr/personal/pdf/KlahrNigam.PsychSci.pdf>.
- Klahr, D. J. L. (2005). Cognitive Research and Elementary Science Instruction: From the Laboratory, to the Classroom, and Back. *Journal of Science Education and Technology*, 14(2), 217-238.
- Kolb. (1971). *Individual learning styles and the learning process*. Cambridge: MIT Press.
- Kolb. (1976). *Learning style inventory : self-scoring test and interpretation booklet*. Boston: McBer.
- Kolb. (1999). *Learning style inventory version 3*. Boston: Hay/McBer.
- Kolb, Boyatzis, E., & Mainemelis. (2002). Learning styles and adaptive flexibility: testing experiential learning theory. P.20. *Human Resources Abstracts*, 37(3), 325-472.
- Kolb, D. A. (1984). *Experiential learning : experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
- Kolb, D. A., Hay Resources, D., & Hay, G. (2007). *The Kolb learning style inventory*. Boston, MA: Hay Resources Direct.
- Lin, J., & Tallman, J. (2006). A Theoretical Framework for Online Inquiry-Based Learning. *Paper presented at the Society for Information Technology & Teacher Education International Conference 2006*, Orlando, Florida, USA.
- Lund, J. L., & Tannehill, D. (2010). *Standards-based physical education curriculum development*. Sudbury, MA: Jones and Bartlett.
- Margianti, E. S. (2001). *Learning environment, mathematics achievement and student attitudes among university computing students in Indonesia*. Retrieved May 1, 2011 from <http://adt.curtin.edu.au/theses/available/adt-WCU20021024.095526>.
- McMahon, J. D. (2007). *Teaching/learning philosophies*. Retrieved July 17, 2011, from <http://www.towson.edu/mcmahon/generic/philosophychart.html>.

- McNaughton, S. (2002). *Meeting of minds*. Wellington, N.Z.: Learning Media.
- Moore, D. W., & Cunningham, J. W. (2006). In D. E. Alvermann, K. A. Hinchman, D. W. Moore, S. F. Phelps, & D. R. Waff (Eds.), *Reconceptualizing the literacies in adolescents' lives* (2nd Ed.). (pp. 129–146), Mahwah, New Jersey: Lawrence Erlbaum.
- Moores, T. T., & Chang, J. C. J. (2009). Self-efficacy, overconfidence, and the negative effect on subsequent performance: A field study. *Information and Management*, 46(2), 69-76.
- Nikolova, N. (2007). *The client-consultant relationship in professional business service firms*. Klner Wiss.-Verl., Kln.
- Nunes, M., & McPherson, M. (2003). Constructivism vs. Objectivism: Where Is Difference for Designers of e-Learning Environments? *Proceedings of the Third IEEE International Conference on Advanced Learning Technologies*, Los Alamitos: IEEE Computer Society, 496-500.
- Palloff, R. M., Pratt, K., & Palloff, R. M. (2007). *Building online learning communities: effective strategies for the virtual classroom*. San Francisco, CA: Jossey-Bass.
- Palmer, S. (2007). An evaluation of streaming digital video resources in on- and off-campus engineering management education. *Computers & Education*, 49(2), 297.
- Paurelle, S. (2003). *E-learning and constructivism: briefing paper, learning and teaching enhancement unit (LTEU)*. Retrieved May 5, 2011, from <http://www.canterbury.ac.uk/support/learning-teaching-enhancement-unit/publications/FLT-briefing-notes/e-learning-and-constructivism.pdf>.
- Piaget, J. (1989). *Studi Sociologici*: Franco Angeli, Milan.
- Reihlen, M., & Apel, B. A. (2007). Internationalization of professional service firms as learning a constructivist approach. *International Journal of Service Industry Management*, 18(2), 140-151.
- Robertson, M., Scarbrough, H., & Swan, J. (2003). Knowledge Creation in Professional Service Firms: Institutional Effects. *Organization Studies*, 24(6), 831-857.
- Ruble, T. L., & Stout, D. E. (1993). Learning Styles and End-User Training: An Unwarranted Leap of Faith. *MIS Quarterly*, 17(1), 115-118.
- Schunk, D. H. (2000). *Learning theories : an educational perspective*. Upper Saddle River, N.J.: Merrill.
- Schunk, D. H. (2008). *Learning theories : an educational perspective*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Scott, R., & Fisher, D. (2001). Science learning environments at the elementary school level in Brunei Darussalam. *Paper presented at the annual meeting of the National Association for Research in Science teaching, St Louis*.
- Selvi, K. (2007). Future Aims of Science Curriculum for Primary School. *Cultura. International Journal of Philosophy of Culture and Axiology*, 8, 76-83.
- Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-Grade Students' Epistemologies of Science: The Impact of School Science Experiences on Epistemological Development. *Cognition and Instruction*, 18(3), 349-422.
- Stone, J. E., & Clements, A. (1998). Research and innovation: Let the buyer beware. In R. Spillane & P. Regnier (Eds.), *The superintendent of the future: strategy and action for achieving academic excellence* (pp. 59-97). Gaithersburg, MD: Aspen.
- Taylor, J. A., & Cox, B. D. (1997). Microgenetic analysis of group-based solution of complex two-step mathematical word problems by fourth graders. *Journal of the Learning Sciences* 6(2), 183–226.
- Taylor, P., & Fraser, B. (1991). Development of an instrument for assessing constructivist learning environments *Paper presented at the annual meeting of the American Educational Research Association*, New York.
- Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring Constructivist Classroom Learning Environments. *International Journal of Educational Research*, 27(4), 293.
- Taylor, P. C., Fraser, B. J., & White, L. R. (1994). CLES: An instrument for monitoring the development of constructivist learning environments. *Paper presented at the American Educational Research Association, New Orleans*.
- Taylor, P. C. S., Fraser, B. J., & White, L. R. (1994). A classroom environment questionnaire for science educators interested in the constructivist reform of school science. *Paper presented at the In A paper presented at the annual meeting of the National Association for Research in Science Teaching*, Anaheim, CA.
- Turner, J. C., Meyer, D. K., Cox, K. E., Logan, C., DiCintio, M., & Thomas, C. T. (1998). Articles - Creating Contexts for Involvement in Mathematics. *Journal of Educational Psychology*, 90(4), 730.

- Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, E. M., Kang, Y., et al. (2002). Articles - The Classroom Environment and Students' Reports of Avoidance Strategies in Mathematics: A Multimethod Study. *Journal of Educational Psychology*, 94(1), 88.
- van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning. *Educational Psychologist*, 38(1), 5-13.
- von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to Learn and Learning to Argue: Case Studies of How Students' Argumentation Relates to Their Scientific Knowledge. *Journal of Research in Science Teaching*, 45(1), 101-131.
- von Glaserfeld, E. (1989). *An Exposition of Constructivism: Why Some Like It Radical*. Retrieved May 1, 2011 from <http://www.oikos.org/constructivism.htm>.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard.
- Werr, A., & Stjernberg, T. (2003). Exploring Management Consulting Firms as Knowledge Systems. *Organization Studies*, 24, 881-908.
- Wertsch, J. V. (1991). *Voices of the mind: a sociocultural approach to mediated action*. Cambridge, MA: Harvard.
- Windschitl, M. (2002). Framing Constructivism in Practice as the Negotiation of Dilemmas: An Analysis of the Conceptual, Pedagogical, Cultural, and Political Challenges Facing Teachers. *Review of Educational Research*, 72(2), 131.
- Woltz, D. J. (2003). Implicit Cognitive Processes as Aptitudes for Learning. *Educational Psychologist*, 38(2), 95-104.
- Woo, M. A., & Kimmick, J. V. (2000). Comparison of Internet versus Lecture Instructional Methods for Teaching Nursing Research. *Journal of Professional Nursing*, 16(3), 132-139.
- Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education*, 10(1), 15-25.
- Yeo, J., & Tan, S. C. (2008). Constructive use of authoritative sources in science meaning making. *Paper presented at the international conference on learning sciences*, Utrecht, The Netherland.
- Yore, L. D., Anderson, J. O., & Shymansky, J. A. (2005). Sensing the Impact of Elementary School Science Reform: A Study of Stakeholder Perceptions of Implementation, Constructivist Strategies, and School-Home Collaboration. *Journal of Science Teacher Education*, 16(1), 65-88.
- Young, L. E., & Maxwell, B. (2007). Teaching nursing: theories and concepts. In Young, L. E., & Paterson, B. L. (Eds.), *Teaching Nursing: Developing a Student-Centered Learning Environment* (pp. 8-19), Philadelphia: Lippincott Williams & Wilkins.
- Zajda, J. (2004). Cultural Transferability and Cross-Cultural Textbook Development. *Education and Society*, 22(1), 83-96.

Ubiquitous English Learning System with Dynamic Personalized Guidance of Learning Portfolio

Ting-Ting Wu¹, Tien-Wen Sung^{2,3}, Yueh-Min Huang^{1,4*}, Chu-Sing Yang² and Jin-Tan Yang⁵

¹Department of Engineering Science, National Cheng Kung University, Taiwan // ²Institute of Computer and Communication Engineering, Department of Electrical Engineering, National Cheng Kung University, Taiwan //

³Department of Applied Information, Hsing Kuo University, Taiwan // ⁴Department of Applied Geoinformatics, Chia Nan University of Pharmacy and Science, Taiwan // ⁵Department of Digital Media Design, Ming Chuan University, Taiwan // n9897104@mail.ncku.edu.tw // kevin@mail.hku.edu.tw // huang@mail.ncku.edu.tw // csyang@mail.ee.ncku.edu.tw // yangdav@gmail.com

*Corresponding Author

ABSTRACT

Situated learning has been recognized as an effective approach in enhancing learning impressions and experiences for students. Can we take advantage of situated learning in helping students who are not English native speakers to read English articles more effective? Can the effectiveness of situated learning be further promoted by individual portfolio? This paper proposes a situated and reading-based English learning system that integrates a reading guidance mechanism into the development of an English learning environment. To facilitate reading, the system offers not only translation of the vocabulary but also the powerful functions such as translation, pronunciation and explanations of sentences, paragraphs and articles. More importantly, the guidance mechanism provides personalized reading suggestions for increasing learning performance according to the assessment of defined dynamic parameters. The assessment refers to a detailed record of the learner behavior, kept as a Learning Portfolio by the system, for improving the accuracy of reading guidance. The experiments show that learners who adopted a situated and reading-based English learning system exhibited higher quality performance than those who adopted conventional learning systems. And, the performance of learners who utilized the learning support system with a reading guidance mechanism was of higher quality than that of learners who merely utilized a simple situated learning system.

Keywords

English reading, Situated learning, Reading guidance, Learning portfolio, Ubiquitous learning

Introduction

Languages are an important means of interpersonal communications. The English language is, in particular, an important one in the international community (Spolsky & Shohamy, 1999). In non-English-speaking countries, the improvement of English proficiencies of nationals has become an important policy to enhance national competitiveness. In the meantime, the rapid advances of information technologies have significantly changed the way people acquire, apply and disseminate information. The advances have also brought magnificent changes and progress to teaching methods in the application of education. The teaching environments that combined computer-aided learning systems with teaching strategies have become an integral part of diversified teaching techniques (Lin & Hsieh, 2001). Furthermore, the growing popularity of wireless technologies and mobile devices have contributed to the development of mobile learning (Chinnery, 2006; Jeng et al., 2010; Klopfer, Squire, & Jenkins, 2002; Soloway et al., 2001). Accordingly, the teaching support system that combined with information and networking technologies to assist the learning of foreign languages has become an important research and development issue (Collins, 2005; Shih, 2005; Sun, Huang & Liu, 2011).

Ubiquitous learning as an educational concept has gradually gaining momentum (Huang et al., in press; Huang, Huang & Hsieh, 2008; Hwang, Tsai & Yang, 2008;). It is being applied to the development of learning support in English-language learning. For example, Ogata and Yano (2004a) proposed the TANGO learning system that allows students to move actual objects in the learning environment according to instructions to enhance the learners' understanding of English vocabulary related to the objects they moved. Follow-up studies combined cooperative learning pedagogy with knowledge maps in a wireless learning environment to help students build English vocabulary (El-Bishouty, Ogata & Yano, 2007; Ogata and Yano, 2004b). Petersen & Markiewicz (2008, 2009) developed PALLAS, a personalized mobile English learning system. The system offers study materials based on the needs and preferences of students to facilitate a personalized learning experience. Chen & Li (2009) established a

ubiquitous learning system that provides suitable vocabulary to the learners based on their vocabulary capacities and wireless base station positioning to enhance their vocabulary levels. Many of these studies focused on the learning and familiarization of English vocabulary as the main goal of the English learning support system. However, pure memorization of English words provides limited help to the improvement of English proficiencies. Students are not all that familiar with when to use what English words or the composition of English sentences. Sometimes, even if students know all the superficial meanings of individual words, they get lost to the implication of the whole sentences. Problems of this kind indirectly prohibit the reading, writing and speaking abilities of students.

To resolve the problems mentioned above, the training in reading comprehension of English-language articles can be applied. Reading is a process where the reader and the reading material establish a meaningful connection in an active and smooth manner (Neil, 1991). Reading ability contains a highly complex cognitive process. The stronger the reading ability, the more able a person is to gather, understand and determine information in order to achieve personal goals and develop potentials (Nunan, 1999). Learning to read in English is the acquisition of an English vocabulary and the establishment of a linguistic capability to understand sentence patterns and syntaxes from reading materials (Hirvela, 2004). It is an important secret of learning any second foreign that one should not only be able to recognize words in articles, but also able to construct the applications and implications of words from the contexts (Day & Bamford, 1998; Gehard, 1996). Fasting and Lyster (2005) found that in the process of teaching experiments, computer-aided reading helps to develop linguistic skills. Some studies indicated that the appropriate use of computer-aided reading can effectively assist teachers to improve the reading proficiency of students in primary and secondary schools (Lynch, Fawcett & Nicolson, 2000; Nicolson, Fawcett & Nicolson, 2000). Some researches indicated that the combination of technological environments and good guidance and reading instructions enhance the linguistic capabilities and enhance reading comprehension of students (Dreyer & Nel, 2003; Stepp-Greany, 2002).

Besides, situated learning involves learners actively participating in a virtual environment and acquiring knowledge through interaction with people, events, and objects in the environment.(Brown et al., 1989).Situated learning (Lave & Wenger, 1991) theory emphasizes a learning situation which stimulates the intrinsic and extrinsic learner motivations, encouraging learners to introduce abstract ideas into concrete thinking, perception, knowledge, and cognition; the learners interact with their environment (Clancey, 1995) and, in doing so, obtain knowledge through context. Palmer & Hornby and other language scholars have proposed situational language teaching theories, stressing that language learning is natural to learners. For practical vocabulary, language should be taught in a context and applied to a particular situation (Hornby, 1950).

Based on the abovementioned advantages of English learning upon reading and situated learning, this paper first outlines the reading-oriented situated learning model. Moreover, because that most previous English learning support tools or systems focus on learning vocabulary, students relatively lack for the ability to infer, analyze, judge, organize, or apply their vocabulary. Even if learners may be proficient in memorization, they may not be able to connect words or interpret sentences. Therefore, the learning system proposed in this paper provides students with the learning support in vocabulary, phrases, sentences, paragraphs and articles. With the situated teaching as the foundation, this paper constructs a ubiquitous English-reading learning system that combines RFID location-aware technologies, uses learners' positions as the situational information, and provides situational English articles that are highly relevant to the specific location and surroundings as learning materials.

Furthermore, when a learning system provides a non-linear and highly autonomous learning model, although students can select contents based on their own decisions and preferences (Duffy, Lowyck & Jonassen, 1993), beginners or passive learners may experience learning disorientation and cognitive overloads (Lin & Gayle, 1996). This affects the thinking and organization of the learning contents, the internalization of knowledge and the effectiveness of learning (Nelson , 1992; Stanton, Correia, & Dias, 2000). The addition of a guidance and support system can mitigate the negative effects (Brusilovsky, 1999; Chen & Hsu, 2008; Hartley & Sleeman, 1973; Maes, 1994). Therefore, the second focus of this paper is to provide appropriate and personalized guidance mechanism for English reading. The guidance mechanism fully utilizes the data of learning portfolio to assess the dynamic parameters defined in the system for suggesting more appropriate reading materials to students in the appropriate sequence.

Ubiquitous English-Reading Learning System

Before doing this experiment, this learning system has been conducted the acceptance and satisfaction evaluation (Wu, Sung, Huang, & Yang, 2010). Based on the result of evaluation data, the learning system has done a minor modification to make it more user-friendly.

Campus Situations for English Learning

This study designs eight campus situations by referring to a university in southern Taiwan. The reading materials offered by the system for students to read and learn are all highly relevant to the student locations and surroundings. The situations and the corresponding materials are as follows:

1. Restaurants and shops: dining, food, order, mall, shopping, etc.
2. Dormitory: daily life, indoor entertainment, etc.
3. Softball field: baseball, softball, Major League, Chinese Professional Baseball League, etc.
4. Library: query and checkout services, book and magazine publications, guided reading, etc.
5. Computer classroom: information, Internet, digital technology, etc.
6. Gym: sports and exercise, indoor competitive ball sports, workout, etc.
7. Teaching complex: teaching activities, tests, class affairs, teachers and students, etc.
8. Parking lot: cars, motorcycles, transportation, etc.

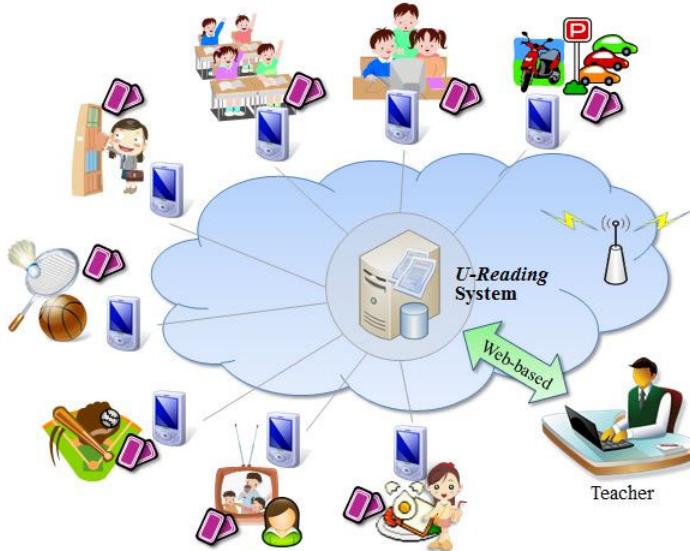


Figure 1. System architecture of U-Reading system

System Architecture

Figure 1 shows the system architecture of this learning system. It is constructed in an IEEE 802.11 WiFi network in a campus. Active RFID tags are installed in the eight situational locations throughout the campus. Students carry with them the mobile learning devices, such as PDAs or smartphones, installed with RFID readers and a built-in WiFi adaptor. Data can be transmitted via wireless networks and RFID technology provides location awareness. Once a mobile learning device senses the location of the student via RFID, the location information will be sent to the U-Reading System which includes Dictionary Server, Learning Material Database, and Portfolio Database. According to the reading guidance mechanism, from the Learning Material Database, the system selects the reading materials highly relevant to the situation and suitable for the students. The student can learn the materials with the supporting functions, e.g. Chinese translation of the vocabulary, phrases, sentences, paragraphs and articles. The audio function allows the students to listen to the pronunciation or recitation. All these features aim to assist learning of the reading

materials. The translation and pronunciation of words and phrases is accessed from Dictionary Server; whilst the translation and pronunciation of sentences, paragraphs and articles are accessed from Learning Material Database. The reading materials are sourced from digital versions of English-learning magazines, audio books, Chinese-English news websites and in-house compilations. Since the system provides functions more than just a directory of English words, it is able to assist the student to understand better the meanings and structures of sentences and the application of words. The learning behavior of each student is collected and recorded in Portfolio Database. Relevant statistics of learning portfolio are calculated and fed back to the Reading Guidance Mechanism as decision-making parameters. Teachers can also access the statistics from the Web-based teachers' interface in order to adjust their teaching strategies and materials accordingly.

Learning Functions and Interface

RFID tags are installed at the appropriate locations throughout the campus and an RFID reader is equipped into the mobile learning device. RFID SDKs support relevant APIs so that the system programs can write, detect and access situational codes with RFID tag memories.

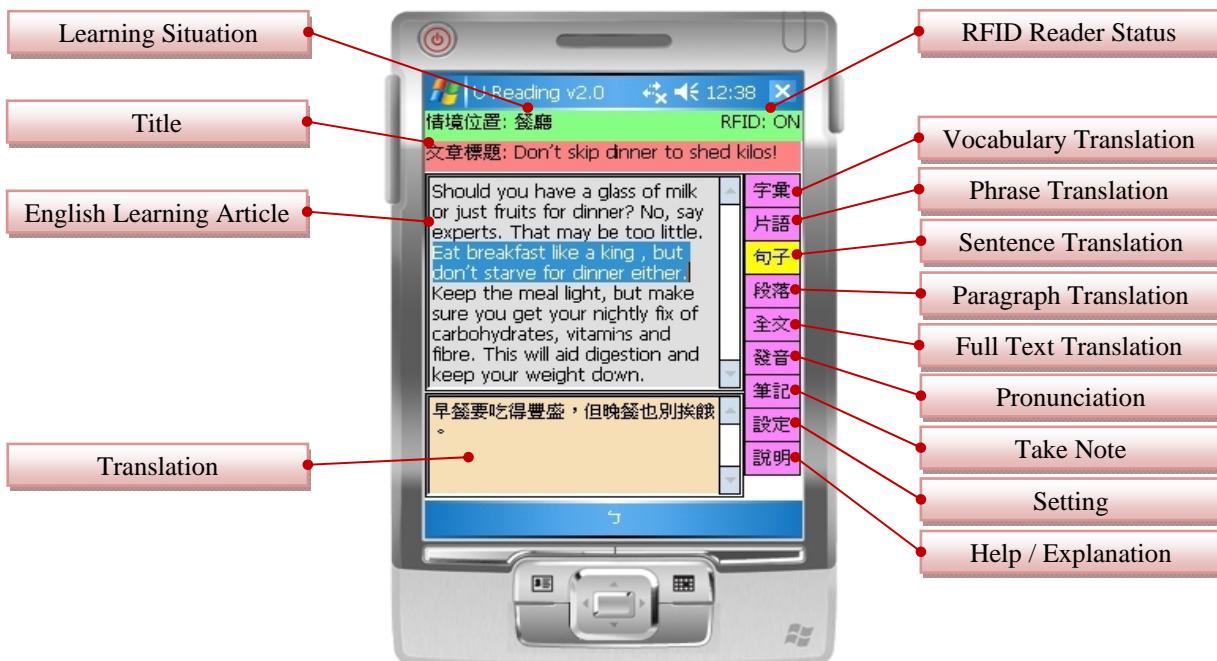


Figure 2. Learning function interface

The English-reading learning interface and the support functions are shown in Figure 2. Once the RFID reader is activated, it can detect and display the location of the user. The system can immediately select suitable reading articles and displayed on the mobile device from the database according to the Reading Guidance Mechanism. Users can inquire for the translation of words, phrases, sentences, paragraphs or whole articles. The inquired results are shown in the lower part of the screen. Users can select the pronunciation or recitation function to improve their speaking and listening. The notebook function logs the inquiries and results to facilitate off-line reviews afterwards. The figure below demonstrates the situation of “restaurant”, as one of the eight campus situations.

Back-end Functions for Teachers

The back-end system provides basic features such as the establishment and maintenance of students' data and teaching materials (Figure 3). The learning behavior is also logged into the back-end database. Therefore, the system also provides Web-based inquiry functions to teachers for all the statistics of the learning portfolio. The statistics

include the number of times and rankings of the usage rates of situations, reading materials, the inquiries of words, sentences, paragraphs and articles for translation and pronunciation. Such information helps teachers to better understand the activities of students so that they can adjust teaching materials or strategies accordingly.



Figure 3. Back-end interface for teachers

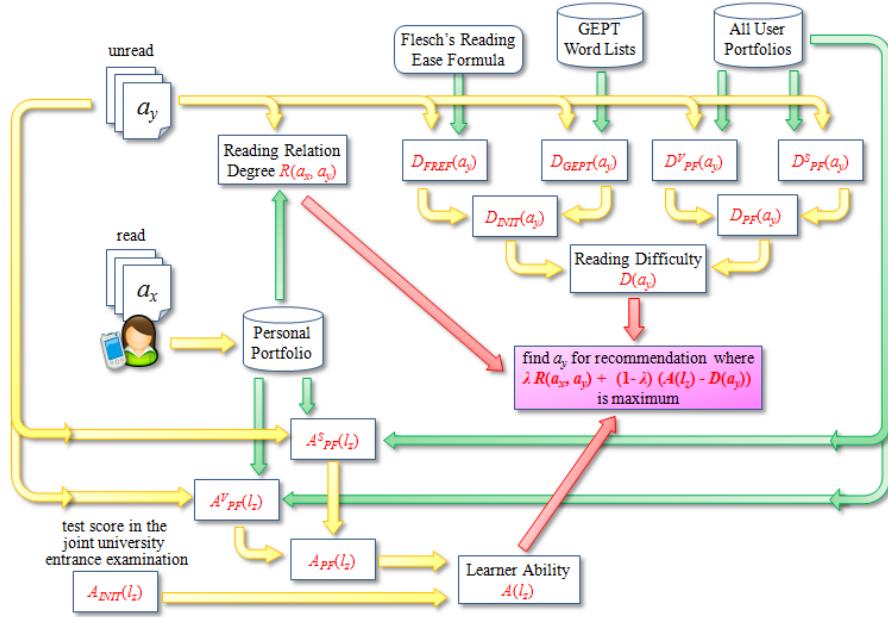


Figure 4. Flow chart of guidance parameter calculation and decision making

Reading Guidance Mechanism

The RFID technology senses the locations and corresponding situations of students to provide situational reading materials. Meanwhile, the system also offers personalized guidance to students to enhance learning effectiveness.

The reading guidance mechanism takes into account the reading relation degree, reading difficulty, and learner ability with regards to the learning portfolio. Learning portfolio can serve as a reference material for the calculation of reading article suggestion (Drachsler et al., 2009). The guidance mechanism dynamically calculated guidance parameters to improve the accuracy of the suggested reading materials and sequence. Figure 4 shows the flow of guidance parameter calculation and decision making. The guidance parameters and decision making are described in detail in the following sections.

Assessment of Reading Relation Degree

The system proposed by Chen (2008) uses the linguistic terms in course description to compute the Concept Relation Degrees for the sequential arrangement of learning courses. This paper defined a different item, Reading Relation Degree, to enhance the relation and proper sequence between suggested reading articles. The system refers to the vocabulary inquired by the student for the previous article in the specific situation to calculate Reading Relation Degree of each unread article in the situation, in order to recommend an article that contains the same vocabulary as much as possible as the previous one. The re-appearance of the same words further familiarizes the students and enhances their understanding of the meanings and usages for these words in different contexts. The Reading Relation Degree for articles a_x and a_y is defined as $R(a_x, a_y)$ and expressed in Eq. (1). The article a_x is the previous article that the student read. The collection of the vocabulary words inquired by the student for that article is $U_x = \{ u_i \mid i \in [1, m_x] \}$, whilst $V_y = \{ v_i \mid i \in [1, n_y] \}$ is the collection of the vocabulary words appear in the article a_y . $W_y = \langle w_1, w_2, w_3, \dots, w_{n_y} \rangle$ is the composition of words in the article a_y . The first part after the equality sign is the percentage of U_x appears in the article a_y . ω denotes the weighting. The second part takes into consideration the reappearance of the same words in order to calculate the percentage of W_y belongs to U_x .

$$R(a_x, a_y) = \omega \sum_{i=1}^{m_x} \frac{f(u_i)}{m_x} + (1 - \omega) \sum_{i=1}^{n_y} \frac{g(w_i)}{h_y} \quad (1)$$

where

$$f(u_i) = \begin{cases} 1 & , u_i \in V_y \\ 0 & , \text{otherwise} \end{cases} \quad (2)$$

$$g(w_i) = \begin{cases} 1 & , w_i \in U_x \\ 0 & , \text{otherwise} \end{cases} \quad (3)$$

$R(a_x, a_y)$ is between 0 and 1. The higher the value, the higher the reading relation degree is.

Assessment of Reading Difficulty

To assess the level of difficulty of reading materials, Chen & Hsu (2008) referred to Flesch Reading Ease Readability Formula (Flesch, 1948) and the vocabulary classification of the GEPT. The assessment considers the differing judgment of English native speakers and non-English native speakers concerning the level of difficulty of reading materials. The benchmark is based on the average standard of the general public. As the target users of this system are university students in Taiwan in the contexts of campus life, this paper refers to the average English proficiencies of students as one of the major criteria in the assessment of the difficulty for reading materials. In Taiwan, all colleges and universities accept enrollments based on the results of joint-entrance examinations. On average, the quality of students is different from one school to another. Therefore, in addition to Flesch's Formula and GEPT, this paper further takes into account the learning portfolio of words and sentences inquired by all the students in order to make real-time dynamic calculations and adjustments for reading difficulty of articles. The Reading Difficulty of the article a_x is defined as Eq. (4). $D_{INIT}(a_x)$ is calculated based on Flesch Reading Ease Readability Formula and the vocabulary classification of the GEPT. This value (between 0 and 1) is used as the

initial value for the predetermined reading difficulty of a_x . $D_{PF}(a_x)$ is the reading difficulty dynamically calculated on the basis of the statistics in the learning portfolio, with α as the weighting.

$$D(a_x) = \alpha \cdot D_{INIT}(a_x) + (1 - \alpha) \cdot D_{PF}(a_x) \quad (4)$$

The Flesch's formula is as follows:

$$RE = 206.835 - (1.015 \times ASL) - (84.6 \times ASW) \quad (5)$$

where

ASL = Average Sentence Length

ASW = Average number of Syllables per Word

The value of RE is between 0 and 100. The higher the value is, the lower the level of difficulty. This paper converts this value into a value between 0 and 1 by Eq. (4), and the higher the value, the more difficult the material is.

$$\begin{aligned} D_{FREF}(a_x) &= 1 - RE(a_x)/100 \\ &= 0.01015 \times ASL(a_x) + 0.846 \times ASW(a_x) - 1.06835 \end{aligned} \quad (6)$$

The GEPT collects more than 8,000 English words and classified these words into Elementary, Intermediate and High-Intermediate. This paper calculates $D_{GEPT}(a_x)$, the difficulty indicator based on the GEPT, and the value is between 0 and 1. A high value means a high difficulty.

$$D_{GEPT}(a_x) = \frac{\frac{1}{6}C_E(a_x) + \frac{3}{6}C_I(a_x) + \frac{5}{6}C_H(a_x)}{C_E(a_x) + C_I(a_x) + C_H(a_x)} \quad (7)$$

$$D_{INIT}(a_x) = \frac{D_{FREF}(a_x) + D_{GEPT}(a_x)}{2} \quad (8)$$

where, $C_E(a_x)$, $C_I(a_x)$ and $C_H(a_x)$ denote the number of words of Elementary, Intermediate and High-Intermediate in article a_x , respectively.

$D_{PF}(a_x)$ is the indicator to the difficulty of reading materials calculated on the basis of learning portfolio consists of the assessment of the vocabulary and sentences. The criteria contain the frequency of the translation inquiries and the capabilities of the inquirer. The calculations are as follows:

$$F_V(a_x) = \sum_{v \in V_x} \frac{q(v)}{p(v) \cdot |V_x|} \quad (9)$$

$$L_V(a_x) = \sum_{v \in V_x} \sum_{i=1}^{q(v)} \frac{c_i(v)}{q(v) \cdot |V_x|} \quad (10)$$

where, $F_V(a_x)$ and $L_V(a_x)$ denote the average frequency of the translation inquiries and average capability of inquirers for the vocabulary appears in the article a_x ; $q(v)$ denotes the total number of times for the inquiries made by all the students on word v ; $p(v)$ denotes the total number of times the word v appears throughout the articles read by all the students; $c_i(v)$ indicates the capability of the learner when the word v is inquired for the i -th time. V_x is the collection of the vocabulary words that appear in the article a_x . Therefore, the indicator to the difficulty for the vocabulary in the article a_x based on the learning portfolio is as follows:

$$\begin{aligned}
D_{PF}^V(a_x) &= \frac{F_V(a_x) + L_V(a_x)}{2} \\
&= \frac{1}{2} \left(\sum_{v \in V_x} \frac{q(v)}{p(v) \cdot |V_x|} + \sum_{v \in V_x} \sum_{i=1}^{q(v)} \frac{c_i(v)}{q(v) \cdot |V_x|} \right)
\end{aligned} \tag{11}$$

Similarly, $q(s)$ is the total number of times that sentence s is inquired by all the students; $p(s)$ denotes the total number of times for sentence s to appear in all the articles read by all the students; $c_i(s)$ is the capability of the learner when the sentence s is inquired for the i -th time; S_x is the collection of the sentences that appear in the article a_x . Therefore, the indicator to the difficulty for the sentences in the article a_x based on the average frequency of the sentence inquiries $F_S(a_x)$, the average capability of the inquirers $L_S(a_x)$ and the learning portfolio is as follows::

$$F_S(a_x) = \sum_{s \in S_x} \frac{q(s)}{p(s) \cdot |S_x|} \tag{12}$$

$$L_S(a_x) = \sum_{s \in S_x} \sum_{i=1}^{q(s)} \frac{c_i(s)}{q(s) \cdot |S_x|} \tag{13}$$

$$\begin{aligned}
D_{PF}^S(a_x) &= \frac{F_S(a_x) + L_S(a_x)}{2} \\
&= \frac{1}{2} \left(\sum_{s \in S_x} \frac{q(s)}{p(s) \cdot |S_x|} + \sum_{s \in S_x} \sum_{i=1}^{q(s)} \frac{c_i(s)}{q(s) \cdot |S_x|} \right)
\end{aligned} \tag{14}$$

$D_{PF}(a_x)$ represents the average responses of a specific group of students to the difficulty of reading materials. The incorporation of this value can enhance the calculation accuracy of the reading difficulty for the reading materials.

$$D_{PF}(a_x) = \frac{D_{PF}^V(a_x) + D_{PF}^S(a_x)}{2} \tag{15}$$

Assessment of Learner Ability

The reading guidance mechanism takes into account the learner ability for offering reading materials. The assessment of learner ability in this paper consists of two parts. The first part is $A_{INIT}(l_y)$, the test score of the student l_y in the subject of English in the joint university entrance examination. This is the initial value for the ability of the student. The second part is the indication value $A_{PF}(l_y)$ calculated real-time and dynamically according to the records in the learning portfolio. Below is the equation for the Learner Ability, i.e. $A(l_y)$, of the student l_y . The value is between 0 and 1. A high value means a strong ability. β denotes the weighting.

$$A(l_y) = \beta \cdot A_{INIT}(l_y) + (1 - \beta) \cdot A_{PF}(l_y) \tag{16}$$

The calculation of $A_{PF}(l_y)$ includes the past performance of the student when facing vocabulary and sentences.

$$A_{PF}(l_y) = \frac{A_{PF}^V(l_y) + A_{PF}^S(l_y)}{2} \tag{17}$$

The following Eq. (18) expresses the past performance of the student when facing vocabulary. One of the indicators is $P_V(l_y)$, which is the percentage of the absence of vocabulary inquiries by student l_y , among all the words that appear in the reading materials previously read by that student, which indicates the performance history of the student l_y . Another indicator is $Q_V(l_y)$, which is the percentage of the inquiries made by all the other students, among all the words inquired by the student l_y , which is a comparison of the ability of the student l_y against the ability of all the other students.

$$A_{PF}^V(l_y) = \frac{P_V(l_y) + Q_V(l_y)}{2} \quad (18)$$

$$P_V(l_y) = \sum_{a_x \in E'} \left(1 - \frac{h_V(a_x)}{|V_x| \cdot |E'|} \right) \quad (19)$$

$$Q_V(l_y) = \sum_{a_x \in E'} \sum_{v \in U_x} \frac{r_V(v)}{|U_x| \cdot |E'|} \quad (20)$$

where, E' denotes the collection of all the articles read by the student l_y ; $h_V(a_x)$ is the number of words inquired by the student for the article a_x ; V_x is the collection of the vocabulary words that appear in the article a_x ; U_x is the collection of the vocabulary words inquired by the student l_y for the article a_x ; $r_V(v)$ is the percentage of the inquiries made by all the other students for the word v in the articles that contain the word v (the number of inquirers divided by the number of readers).

$$A_{PF}^V(l_y) = \frac{1}{2} \left(\sum_{a_x \in E'} \left(1 - \frac{h_V(a_x)}{|V_x| \cdot |E'|} \right) + \sum_{a_x \in E'} \sum_{v \in U_x} \frac{r_V(v)}{|U_x| \cdot |E'|} \right) \quad (21)$$

Similarly, the ability concerning the comprehension of the sentences is expressed as following Eq. (22) and Eq. (23). The percentage of the absence of the inquiries by the student l_y is defined as $P_S(l_y)$. $Q_S(l_y)$ is the percentage of the inquiries made by other students for the sentences also inquired by the student l_y .

$$A_{PF}^S(l_y) = \frac{P_S(l_y) + Q_S(l_y)}{2} \quad (22)$$

$$= \frac{1}{2} \left(\sum_{a_x \in E'} \left(1 - \frac{h_S(a_x)}{|S_x| \cdot |E'|} \right) + \sum_{a_x \in E'} \sum_{s \in T_x} \frac{r_S(s)}{|T_x| \cdot |E'|} \right) \quad (23)$$

where, $h_S(a_x)$ is the number of sentences inquired by the student l_y for the article a_x ; S_x is the collection of the sentences that appear in the article a_x ; T_x is the collection of the sentences inquired by the student l_y for the article a_x ; $r_S(s)$ is the percentage of the inquiries made by all the other students for the sentence s in the articles that contain the sentences (the number of inquirers divided by the number of readers).

Recommended Reading Materials

The guidance is about providing the next best option for learners based on their needs (Eklund & Sinclair, 2000). As shown in Eq. (24), this paper selects the article with the highest assessment value, i.e. a_{fit} , for the learner l_z based on the evaluation of the Reading Relation Degree, Reading Difficulty, Learner Ability.

$$\begin{aligned}
a_{fit} &= \arg \max_{a_y \in E-E'} K_{fitness}(a_y) \\
&= \arg \max_{a_y \in E-E'} (\lambda \cdot R(a_x, a_y) + (1-\lambda)(A(l_z) - D(a_y)))
\end{aligned} \tag{24}$$

where, E is the collection of articles; E' is the collection of the articles read completely; $K_{fitness}(\cdot)$ is the evaluation function of the suitability of the reading materials, λ is the weighting of the evaluation parameter.

Experiment

Design

This study evaluates the students of three classes in the freshman year of a university. Each class is defined as one group for the evaluation. Three groups, which are group P (Paper), group R (Reading), and group G (Guidance), are established. The students of group P learn from paper-based materials and do not access any support of the learning system. The teaching is in the form of traditional lectures in the classroom. Both group R and group G use the Ubiquitous English-Reading Learning System, with the only difference that group G has additional assistance from the personalized reading guidance mechanism. The experiment was designed to compare the learning effectiveness of these three groups and to evaluate whether the system can enhance learning motives, interests, and a sense of participation. The benchmark was the reading materials centering on the situation of restaurants and shops. The goal was to assess the differences in learning effectiveness of the traditional classroom approach, the u-learning system, and the u-learning system combined with the reading guidance mechanism. We attempted to understand if the learning performance of learners with the reading guidance mechanism (group G) proposed by this study was higher than that of other two groups (group R & group P). Did learners who adopted the English learning system (group R & group G) exhibit higher learning performance than those who adopted conventional learning methods (Group P)? All learning content was identical and came from the same Learning Material Database.

Participants

The research participants are the freshman students learning English in a university in Taiwan. There are a total of 113 students from the selected three classes. One class consists of 36 students (22 males and 14 females), another class 39 students (23 males and 16 females), the third class 38 students (22 males and 16 females). The teacher is very interested in the application of technologies into English-language teaching and grasping the learning status of students, thus is cooperative in this experiment.

Procedure

In this experiment, there are 18 weeks each semester for the English-language curriculum, with two classes per week. The 9th and 18th weeks are dedicated to the mid-term and final exams, respectively. From the 1st to the 8th week, traditional lecturers are given with teachers explaining the basis of contexts and grammar. A pre-test is performed in the 9th week, to assess the level of familiarity of each student. From the 10th to the 17th week, the experimental teaching is given. To ensure that both the teacher and students are familiar with the learning flows and the operations of the system devices, an orientation on the learning flows and the device instructions is held in the 10th week. From the 11th week onward, group P studies the reading materials in the conventional manner; whereas group R and group G begin to use the mobile devices equipped with WiFi and RFID components and the proposed Ubiquitous English-Reading Learning System for English reading (Figure 5). In the seven weeks of experimental learning, the teacher arranges a feedback session in the 14th week. Based on the records and statistics shown on the back-end system and the responses from the students, the teacher corrects the behavior bias and cognitive mistakes of students to facilitate further learning. The experimental learning finishes at the end of 17th week. The Learning Material Database at the backend contained roughly 40 articles for each campus situation. Each student was required to read at least 20 articles for the situation of restaurants and shops during the experiment. In the last week, i.e. the 18th week, students take the final written examination (post-test) and fill in the questionnaire concerning the system acceptance.

Limitations

This experiment adopted the learning method that required high autonomy, which means participants were required to exhibit high autonomy and self-discipline during the experiment. Moreover, during the experiment the charge capacity of mobile devices was limited, and learners who were not familiar with the operation of the devices might suffer from crashing the system and difficulty of operation.

In the article recommending guidance mechanism, this paper combined data of individual and group learning portfolios on the decision-making calculation of recommended guidance. The higher the frequency of recorded reading behaviors in the learning portfolios database, the more accurate the results are.



Figure 5. The learning situation of restaurant

Data Collection

Pre-test and Post-test Scores

Both the pre-test and the post-test are based on the questions designed by the teacher. The questions are about the location and surroundings of the situation of restaurants and shops. The teacher designs the questions also by referring to the teaching materials in the Learning Material Database. The questions are in the form of vocabulary explanations, blank fillings, sentence translation and reading comprehension. The test results indicate whether the proposed Ubiquitous English-Reading Learning System can help students to improve their English proficiency and understanding. Assessments are also made to gauge whether the reading guidance mechanism is beneficial to English learning.

Survey on Suitability of Recommended Reading Materials

After a student finishes reading each English-language article, the system automatically presents a questionnaire for the student to fill in replies concerning the level of difficulty and relevance of the article. The analysis aims to evaluate the suitability of the recommended reading materials to the students with and without access to the reading guidance mechanism.

Questionnaire on System Acceptance

The ease of use and practicality of the system will influence the intention of learners in using the system and hence their learning effectiveness (Davis, 1989). This study performs a questionnaire survey on the system acceptance to gather feedbacks for further designs and improvements. The questionnaire is designed based on the questionnaire developed by Davis (1989) and Harwick & Barki (1994). The measurement is based on the Likert 5-point scale. The criterion-related validity and construct validity are ensured with the assistance and correction from two experts. The

Cronbach α coefficient is above 0.7, indicating good consistency and stability. A total of 75 questionnaires are issued. After the elimination of 10 ineffective questionnaires, there are 65 effective questionnaires. The effective recovery rate is 87%.

Results

Assessments of Learning Effectiveness

Within-Group Assessment

This analysis is based on the t-tests of dependent samples of group *P*, group *R* and group *G*. Table 1 summarizes the statistical analysis of the variances in the pre-test and post-test results.

Table 1. The statistical analysis of within-group learning outcome

| Group | Test | N | Mean | Std. Deviation | t | p |
|----------------|-----------|----|-------|----------------|---------|-------|
| group <i>P</i> | Pre-test | 38 | 65.16 | 5.726 | -1.471 | .150 |
| | Post-test | 38 | 66.21 | 4.592 | | |
| group <i>R</i> | Pre-test | 39 | 64.74 | 5.571 | -11.398 | .000* |
| | Post-test | 39 | 80.97 | 6.714 | | |
| group <i>G</i> | Pre-test | 36 | 64.97 | 5.882 | -18.817 | .000* |
| | Post-test | 36 | 84.53 | 5.853 | | |

$\alpha = .05$; * $p < .05$

The analysis indicates that the students in group *P* make very limited progress from the traditional learning approach in classroom. However, the overall learning effectiveness is far inferior to the students in the other groups who use the learning system. The reason should be a lack of effective support, learning fun and motives. In contrast, the students in both group *R* and group *G* show significant improvements in learning effectiveness after they use the learning system. This shows a convenient reading support and the situated learning (of reading what you see) enhances an understanding of the English language and betters the learning outcomes. The students in group *G* with additional access to the reading guidance mechanism report an average of progress scores 3.33 higher than that of group *R*. This indicates that the reading guidance mechanism with multi-facet consideration can provide more appropriate reading materials and sequence dynamically to the students. The reading relevance can enhance the connection of learning content and memory, and the students' ability is matched with the recommended articles, thus can reduce learning obstacles and enhance learning effect.

Between-Group Assessment

Between-Group Assessment included a pre-test and a post-test. In the pre-test, we adopted a one-way ANOVA to discuss the English performance of group *P*, group *R*, and group *G* before the experiment. The results of statistical analysis are shown in Table 2 below.

Table 2. The statistical pre-test analysis between groups

| Group | Mean | N | Std. Deviation | F | p |
|----------------|-------|----|----------------|------|------|
| group <i>P</i> | 65.16 | 38 | 5.726 | | |
| group <i>R</i> | 64.74 | 39 | 5.571 | .051 | .951 |
| group <i>G</i> | 64.74 | 39 | 5.571 | | |

* $p < .05$

The average pre-test scores of group *P*, group *R* and group *G* are 65.16, 64.97 and 64.74, respectively. These scores are similar. The *p* value is greater than 0.05, indicating a lack of significant variances in the learning effectiveness of these three groups. There are no obvious differences in the levels of English-language test results since the university ensures that the levels of students in each class are in a normal distribution.

Table 3 shows the analysis of the post-test results of group *P*, group *R* and group *G* by using ANOVA of independent samples. The purpose is to understand that the learning effectiveness of the group of students in group *P* and group *R* that use the learning system. The analysis also aims to evaluate the learning outcomes of group *R* and group *G* who use the reading guidance mechanism of the system.

Table 3. The statistical post-test analysis between groups

| Group | Mean | N | Std. Deviation | F | p |
|----------------|-------|----|----------------|---------|-------|
| group <i>P</i> | 66.21 | 38 | 4.592 | | |
| group <i>R</i> | 80.97 | 39 | 6.714 | 105.426 | .000* |
| group <i>G</i> | 80.97 | 39 | 6.714 | | |

*p < .05

As shown in Table 3, the value p of statistical analysis is less than .05. The result shows that the scores of the post-test in group *P*, group *R*, and group *G* had significant differences. We further observed that the scores of the post-test in group *P* were 14.76 lower than the scores in group *R*. The scores of the post-test in group *R* were significantly higher than the scores in group *P*. This shows that the learning effectiveness of group *R*, who uses the system, is significantly better than the students who resort to traditional learning approaches. It confirms that the learning supports of the proposed system can effectively enhance learning results, and furthermore, the system can effectively connect the reading materials with situations, thus reducing abstraction. It can effectively help students to internalize the contents and improve their understanding of the English language so that they can use the language better. Additionally, the average score of group *G* is higher than that of group *R*. These analysis results demonstrate that although both groups use the learning system, the group that has access to the reading guidance mechanism benefits more give the personalized guidance based on the reading relevance and level of difficulty of learning materials, and ability of learners. In addition, the standard deviation of group *G* is smaller than that of group *R*, indicating that the system is conducive to closing the gap between the students.

Suitability of Recommended Reading Materials

Table 4 shows the suitability of the recommended reading materials. A questionnaire survey is performed on group *R* and group *G*. The questionnaire is based on Likert 5-point point scale and the scores are calculated with percentages. Regarding the level of difficulties suitable to the capabilities of learners, the survey shows that the average score of group *G* is 4.04, higher than that of group *R* at 3.43, a clear indication of the results of the learning guidance mechanism. About 30% of the group *R* respondents disagree that the level of difficulties is appropriate to their capabilities because there is a lack of reading guidance and a sequence of recommended articles.

Table 4. Suitability of recommended reading materials

| Questions | strongly disagree → strongly agree | | | | | average | |
|---|------------------------------------|-----|-----|-----|-----|---------|------|
| | 1 | 2 | 3 | 4 | 5 | | |
| Level of difficulties suitable to my own capabilities | group <i>R</i> | 2% | 28% | 10% | 45% | 15% | 3.43 |
| | group <i>G</i> | 0% | 5% | 16% | 49% | 30% | 4.04 |
| Sufficient relevance of the reading materials | group <i>R</i> | 11% | 19% | 27% | 25% | 18% | 3.20 |
| | group <i>G</i> | 4% | 10% | 23% | 38% | 25% | 3.70 |

As for the relevance of reading materials, although group *R* does not use the reading guidance mechanism, 43% of the respondents think it is sufficient. The percentage is greater than 30% of the respondents who do not think it is sufficient. The average of the 5-point scale measurement is 3.20, not too poor a number. This is because the reading materials for the same situation are high relevant already. The users in group *G* use the reading guidance mechanism and achieve an average of 0.5 higher than group *R*.

There is an additional remark concerning the suitability of recommended reading materials. The results of both the difficulty and relevance of the recommended reading materials could still be improved if the backend database significantly expands its collection of reading materials.

Assessment of System Acceptance

The questionnaire survey on the system acceptance is performed on group *R* and group *G*, in order to understand the learners' satisfaction and acceptance regarding the learning system. The survey results are summarized in Table 5.

Table 5 shows that the average values of most questionnaire replies are greater than 3.0. This indicates that most learners are positive about this system in their study of the English-language materials. About 77% of the learners believe that the contents are appropriate and can meet their English-learning requirements. As many as 85% of the learners think that such a learning system is beneficial to the overall English-language learning. Meanwhile, the fact that the contents are highly relevant to the surroundings helps to reinforce memorization and learning effectiveness. About 80% of the learners are willing to continue with the system and more than happy to recommend the system to others. The system is generally well-received as its learning mechanism and supporting functions can allow learners to achieve different results. The average values for the responses concerning the appropriate size of the display screen and the appropriate size of fonts were lower than those of other responses. As this learning system is installed on PDAs, inherent limitations to font and screen sizes persist. Meanwhile, it is the first time that the surveyed learners use this new learning method and learning tool, few of them need time to adjust to it. Most of the surveyed learners find this novel way of learning acceptable and believe that this system can be a great option to improve their English learning.

Table 5. The results of learners' system acceptance

| Questions | strongly disagree → strongly agree | | | | | average |
|---|------------------------------------|-----|-----|-----|-----|---------|
| | 1 | 2 | 3 | 4 | 5 | |
| The fonts in the system were of appropriate size. | 14% | 36% | 28% | 13% | 9% | 2.67 |
| The buttons in the system were clear and easy to use. | 0% | 8% | 23% | 48% | 21% | 3.82 |
| The operation guidelines of the system were clear and easy to use. | 7% | 13% | 25% | 35% | 20% | 3.48 |
| The colors in the system were clear and easy to differentiate. | 2% | 9% | 19% | 47% | 23% | 3.80 |
| The browsers in the system were of appropriate size. | 18% | 40% | 26% | 11% | 5% | 2.45 |
| The system messages were easy to understand. | 1% | 8% | 19% | 37% | 35% | 3.97 |
| The learning contents provided in the system were appropriate to the learners' needs. | 0% | 10% | 13% | 45% | 32% | 3.99 |
| I have a stronger impression of the English content I learned through this kind of learning method. | 0% | 7% | 15% | 42% | 36% | 4.07 |
| I like the design of the learning system interface. | 0% | 10% | 20% | 48% | 22% | 3.82 |
| It's easier for me to learn through this learning method than when learning from a teacher | 1% | 6% | 12% | 46% | 35% | 4.08 |
| This kind of learning method helps me to learn anytime I want. | 0% | 10% | 20% | 40% | 30% | 3.90 |
| I am very interested in this kind of English learning system. | 6% | 12% | 18% | 42% | 22% | 3.62 |
| The contents introduced in the system are related to my life. | 0% | 10% | 12% | 45% | 33% | 4.01 |
| This system helps me to learn English. | 2% | 6% | 10% | 46% | 36% | 4.08 |
| I enjoy learning through this system. | 1% | 9% | 10% | 48% | 32% | 4.01 |
| I would be happy to recommend this system to other learners. | 0% | 10% | 12% | 42% | 36% | 4.04 |
| This system helps me improve my English reading abilities. | 1% | 8% | 12% | 45% | 34% | 4.03 |

Discussion and Conclusion

This paper designs a learning system that can effectively enhance the English-proficiency of students in non-English speaking countries by combining wireless technologies, RFID equipment and mobile devices. The system provides comprehensive supporting features to assist the learning of reading materials by making up the inefficiency of vocabulary-centric memorization-oriented learning model. It can truly improve the ability of students in learning from the reading materials (from words, phrases, sentences, paragraphs to whole articles) in terms of organization of sentences, application of words and understanding of meanings and implications. The system is designed based on the theory of reading-based and situated learning to recommend reading materials that highly relevant to the surroundings at the location of the learners in order for learners to enhance their learning effectiveness. To further

enhance learning outcomes, it importantly utilizes the records of learning portfolio to offer reading guidance by taking into account the reading relation degree, difficulty degree and learner ability. All data are dynamically computed in real-time to guide the decision-making indicators and provide progressive reading suggestions to learners in the situated environment. This approach is beneficial to the digestion and internalization of reading contents. A series of evaluations on freshman students in universities show that the results of this system are positive. According to the results of evaluation, learners who had adopted situated English learning system improved significantly in their English performance and increased their interests in learning English. This English learning environment combines the theory of situated learning can enlighten students converting abstract perception into concrete thinking. Moreover, this learning strategy can let students practice applications of words, phrases and sentences reflective to the situations which they are in. Those results further confirmed related theories of situated language teaching. Learners who were provided with reading guidance mechanism in learning system scored significantly higher in the post-test after English lessons than the other two groups did. We could learn from this phenomenon that including learning guidance support mechanism in the learning system and providing adequate reading articles to learners could efficiently reduce learning obstacles and increase learning efficiency. Besides, according to the analysis of questionnaire, most learners believed that articles recommended by the system were appropriate to the proficiency of learners, and they also believed this kind of learning system benefited English learning and satisfied the need of English learning. Currently, the system operates in a campus environment. However, it can surely extend to daily life situations to assist the general public in learning English. Future studies can add Intelligent Agents to the system to provide features such as automatic searches, screening, classification and filing of reading materials to effectively expand the Learning Material Database in order to enhance the richness and diversity of the reading contents. Moreover, we can focus on analyzing more detailed learning portfolio arguments during learning process. Those may include the relevance of the time spent on article reading, the influence on proficiency of learners and difficulty of articles, the possibility of auxiliary voice to improve reading pronunciation and the influence of weight coefficient differences on the formula of guidance mechanism. We expect that this study can provide perspectives and directions worthy of studying for academic research in related fields by subsequent detailed analysis and induction.

Acknowledgments

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 97-2511-S-006-001-MY3, NSC 97-2511-S-218-002-MY3, and NSC 99-2631-S-011 -002.

References

- Brusilovsky, P. (1999). Adaptive and Intelligent Technologies for Web-based Education. *Künstliche Intelligenz*, 4, 19-25.
- Chen, C. M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 5(2), 787-814.
- Chen, C. M., and Li, Y. L. (2009). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18 (4), 341-364.
- Chen, C. M., & Hsu, S. H. (2008). Personalized Intelligent Mobile Learning System for Supporting Effective English Learning. *Educational Technology & Society*, 11 (3), 153-180.
- Chinnery, G. (2006). Emerging Technologies - Going to the MALL: Mobile Assisted Language Learning. *Language Learning & Technology*, 10(1), 9-16.
- Clancey, W. J. (1995). A tutorial on situated learning. *Proceedings of the International Conference on Computers and Education*, AACE, 49-70.
- Collins, T. G. (2005). English Class on the Air: Mobile Language Learning with Cell Phones. *Paper presented at the Fifth IEEE International Conference on Advanced Learning Technologies*, July 5-8, 2005, Kaohsiung, Taiwan.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340.
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Drachsler, H., Hummel, H., van den Berg, B., Eshuis, J., Waterink, W., Nadolski, R., Berlanga, A., Boers, N., & Koper, R. (2009). Effects of the ISIS Recommender System for navigation support in self-organised Learning Networks. *Educational Technology & Society*, 12 (3), 115–126.

Dreyer, C. & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System*, 31, 349-365.

Duffy, T. M., Lowyck, J. & Jonassen, D. H. (1993). *Designing environment for constructive learning*. Heidelberg: Springer.

Eklund, J. and Sinclair, K. (2000). An empirical appraisal of the effectiveness of adaptive interfaces for instructional systems. *Educational Technology & Society*, 3(4), 165-177.

El-Bishouty, M. M., Ogata, H., & Yano, Y. (2007). PERKAM: Personalized Knowledge Awareness Map for Computer Supported Ubiquitous Learning. *Educational Technology & Society*, 10(3), 122-134.

Fasting, R. B., & Lyster, S.-A. H. (2005). The effects of computer technology in assisting the development of literacy in young struggling readers and spellers. *European Journal of Special Needs Education*, 20(1), 21-40.

Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, 32(3), 221-233.

Gehard, J. G. (1996). *Teaching English as a foreign or second language*. Ann Arbor, MI: University of Michigan Press.

Hartley, J. R., & Sleeman, D.H. (1973). Towards more intelligent teaching systems. *International Journal of Man-Machine Studies*, 2, 215-236.

Harwick, J., & Barki, H. (1994). Explaining the role of user participation in information systems use. *Management Science*, 40(4), 440-65.

Hirvela, A. (2004). *Connecting reading & writing*. Ann Arbor, MI: The University of Michigan Press.

Hornby, A. S. (1950). The situational approach in language teaching. A series of three articles in *English Language Teaching*. 4, 98-104, 121-8, 150-6.

Huang, Y. M., Huang, T. C., & Hsieh, M. Y. (2008). Using annotation services in a ubiquitous Jigsaw cooperative learning environment. *Educational Technology & Society*, 11(2), 3-15.

Huang, Y. M., Chiu, P. S., Liu, T. C., & Chen, T. S. (in press). The Design and Implementation of a Meaningful Learning-based Evaluation Method for Ubiquitous Learning. *Computers & Education*, doi:10.1016/j.compedu.2011.05.023.

Hwang, G. J., Tsai, C. C., & Yang, S. J. H. (2008). Criteria, Strategies and Research Issues of Context-Aware Ubiquitous Learning. *Educational Technology & Society*, 11(2), 81-91.

Jeng, Y. L., Wu, T. T., Huang, Y. M., Tan, Q., & Yang, S. J. H. (2010). The add-on impact of mobile applications in learning strategies: a review study. *Educational Technology & Society*, 13(3), 3-11.

Klopfer, E., Squire, K., & Jenkins, H. (2002). Environment detectives: PDAs as a window into a virtual simulated world. *Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education*. Vaxjo, Sweden: IEEE Computer Society, 95-98.

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Lin, B., & Hsieh, C. (2001). Web-based Teaching and Learner Control: a Research Review. *Computers & Education*, 37, 377-386.

Lin, C. H. & Gayle, D. (1996). Effects of linking structure and cognitive style on students' performance and attitude in a computer-based hypertext environment. *Journal of Educational Computing Research*, 15(4), 317-329.

Lynch, L., Fawcett, A. J., & Nicolson, R. I. (2000). Computer-assisted reading intervention in a secondary school: an evaluation study. *British Journal of Educational Technology*, 31(4), 333-348.

Maes, P. (1994), Agents that reduce work and information overload. *Communication of the ACM*, 37(7), 30-40.

Neil, A. (1991). Individual difference in strategy use in second language reading and testing. *The Modern Language Journal*, 75, 460-472.

Nelson, W. (1992). Adaptive hypermedia instructional systems: Possibilities for learner modeling. ERIC Document Reproduction Service N0. ED 347195.

Nicolson, R. I., Fawcett, A. J., & Nicolson, M. K. (2000). Evaluataion of a computer-based reading intervention in infant and junior schools. *Journal of Research in Reading*, 23(2), 194-209.

Nunan, D. (1999). *Second Language Teaching and Learning*. Boston: Heinle & Heinle Publishers

- Ogata, H., & Yano, Y. (2004a). Context-aware support for computer-supported ubiquitous learning. *Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education*, 27–34.
- Ogata, H., & Yano, Y. (2004b). Knowledge Awareness Map for Computer-Supported Ubiquitous Language-Learning. *Proceedings of the WMTE'04*, 19-26.
- Petersen, S. A., & Markiewicz, J.-K. (2008). PALLAS: Personalised Language Learning on Mobile Devices. *Proceedings of Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education*(WMUTE2008), March 23-26, Beijing, China, 52-59.
- Petersen, S. A., & Markiewicz, J.-K. (2009). Personalized and contextualized language learning: Choose when, where and what. *Research and Practice in Technology-Enhanced Learning*, 4(1), 33–60.
- Shih, Y. E. (2005). Language in Action: Applying Mobile Classroom in Foreign Language Learning. *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies*, July 5-8, 2005, Kaohsiung, Taiwan.
- Soloway, E., Norris, C., Blumenfeld, P., Fishman, B., Krajcik, J., & Marx, R. (2001). Log on education: Handheld devices are ready-at-hand. *Communications of the ACM*, 44(6), 15-20.
- Spolsky, B., & Shohamy, E. (1999). *The languages of Israel: Policy, ideology, and practice*. Clevedon, UK: Multilingual Matters.
- Stanton, N., Correia, A. P. & Dias, P. (2000). Efficacy of a map on search, orientation and access behavior in a hypermedia system. *Computer & Education*, 35, 263-279.
- Stepp-Greany, J. (2002). Student Perceptions on Language Learning in a Technological Environment: Implications for the New Millennium. *Language Learning & Technology*, 6(1), 165-180.
- Sun, K.T., Huang, Y.M., & Liu, M.C. (2011). A WordNet-Based Near-Synonyms and Similar-Looking Word Learning System. *Educational Technology & Society*, 14(1), 121-134.
- Wu, T. T., Sung, T. W., Huang, Y. M., & Yang, C. S. (2010). Location Awareness Mobile Situated English Reading Learning System. *Journal of Internet Technology*, 11(7), 923-934.

A New Approach Toward Digital Storytelling: An Activity Focused on Writing Self-efficacy in a Virtual Learning Environment

Yan Xu¹, Hyungsung Park^{2*} and Youngkyun Baek³

¹Department of Educational Technology, Korea National University of Education, South Korea // ²Department of Teacher Education, Kyonggi University, South Korea // ³Department of Educational Technology, Boise State University, USA // xuyan1218@hotmail.com // hyungsung@gmail.com // youngkyunbaek@boisestate.edu

*corresponding author

ABSTRACT

Recently, computer technology and multimedia elements have been developed and integrated into teaching and learning. Entertainment-based learning environments can make learning contents more attractive, and thus can lead to learners' active participation and facilitate learning. A significant amount of research examines using video editing software to create video clips for digital storytelling activities. However, in this study, we suggest performing digital storytelling in virtual worlds with open-ended, edutainment elements, and place more stress on the writing process. The purpose of this study is to examine the effects of writing for digital storytelling on writing self-efficacy and on flow in the virtual reality learning environment known as Second Life. The researchers organized an activity for undergraduate students to create digital stories. Participants were divided into two groups. One group created their digital stories in Second Life. The other group created their digital stories off-line. The two groups' independent sample t-tests were employed to compare writing self-efficacy and flow. The results of the experiment demonstrate that digital storytelling in a virtual learning environment is more effective than digital storytelling off-line. The findings of this study suggest that the technique of digital storytelling can be used effectively in classroom settings to teach writing.

Keywords

Digital storytelling, Virtual learning environment, Second Life, Writing self-efficacy, Virtual reality

Introduction

Everyone has their stories to tell, and every day we hear from others about their experiences in the form of stories. Storytelling is a quintessentially human experience, whether conveying stories about oneself, others, or about the world (McDrury & Alterio, 2003). The stories can be real or fictitious. Storytelling is widely used in classrooms to enrich the learning experience. Storytelling is a natural method of human communication and is prevalent in all aspects of human social interaction. People tend to make better sense of complex ideas, concepts, or information when it occurs via storytelling (Chung, 2006). In this context, Ibanez et al. (2003), Mello (2001), and Sadik (2008) have said that storytelling can be used to enhance a student's higher-order thinking and literacy skills, thereby increasing collaborative learning. Nowadays, with the rapid development of technology, a new version of storytelling is emerging: digital storytelling. Digital storytelling integrated with powerful technology can be an effective tool to enhance teaching and learning. The educational application of digital storytelling is attracting the attention of many educators (Chung, 2006; Lee, 2005; Robin, 2008; Sadik, 2008; Shin & Park, 2008).

Kang et al. (2003), Shin and Park (2008) define "digital storytelling" as storytelling that is conducted using digital technology as the medium or method of expression, in particular using digital media in a computer-network environment. Digital storytelling encompasses these key characteristics: Flexibility, universality, interactivity and community formation (Park & Seo, 2009). Flexibility in digital storytelling refers to the creation of a non-linear story using digital media technology (Gregori-Signes, 2008). Universality means that anyone can become the producer of digital stories, due to the wide availability of computers and easy-to-use software. Interactivity refers to the participation of users in the development of the stories using media characteristics that can be mutually exchanged.

Paul and Fiebich (2002) argue that digital stories can be presented in a variety of formats such as text web pages, a nonlinear interactive website, a digital song, a digital video(Chung, 2006), an online game, or a virtual reality world(Shin & Park, 2008). In particular, virtual reality learning environments have the potential to provide rich and engaging learning experiences for students that include investigation, discovery, and creation (Dickey, 2005; Park & Baek, 2009; Robin, 2008). Students experience 'learning by doing' while making a digital story. They perform multiple tasks within the virtual world as researchers, playwrights, designers, media producers, and educators (Chung, 2006).

Communication seems to be more effective when it occurs through digital storytelling. The student can experience storytelling activity in a virtual-reality learning environment such as Second Life. Second Life allows synchronous communication with others, using avatars to generate a sense of presence. Participants write their stories dynamically through interactive communication when they meet, visit the land, work, and talk to other residents of Second Life.

According to Robin (2007), writing is an essential part of the process of creating digital stories. Various research studies have verified the effectiveness of digital storytelling in improving students' writing skills (Ballast, Stephens, & Radcliffe, 2008; Gakhar & Thompson, 2007). It is, therefore, worth thinking about virtual worlds, which contain various multimedia components, as an ideal medium for digital storytelling. Creating digital stories in an immersive virtual location may attract and motivate people into writing unconsciously. Ultimately, digital storytelling may encourage people who don't like writing or who have no confidence in writing.

Over the past decades, a significant amount of research examined using video editing software (such as Windows Movie Maker) to create video clips for digital storytelling activity. However, in this study, we put more emphasis on the writing process and suggest performing digital storytelling in virtual worlds with open-ended, edutainment elements. The purpose of this study is to examine the effects of digital storytelling on writing self-efficacy and flow in the virtual reality learning environment such as Second Life virtual world.

Literature Review

Benefits and Procedures of Digital Storytelling in virtual worlds

Definitions of digital storytelling differ among researchers, but demonstrate several important similarities and main ideas. The definition from The Digital Storytelling Association (2002) describes digital storytelling as giving expression to old forms of storytelling in a modern way. People deliver their knowledge, values, and wisdom with stories. Emerging technologies make it possible to see such stories on the computer screen. According to Armstrong (2003), digital storytelling is the same as reading and writing stories. However, digital storytelling's defining feature in education is that it shares information through multimedia.

The University of Houston (2009) Instructional Technology Department conducts studies examining the educational uses of digital storytelling. They suggest the following procedures for digital storytelling: In the first stage, the storyteller defines the parameters of the story. The learner should select a topic for the digital story. Next, he should search for image resources for the story (pictures, drawings, photographs, maps, and/or charts), audio resources (music, speeches, interviews, and/or sound effects), and informational content (perhaps from web sites, word processed documents, or PowerPoint slides). When the student has gathered all of his resources, he should begin thinking about the purpose of the story. Is the purpose to inform, convince, provoke, or question?

In the second stage, the storyteller organizes and selects specific audio, images, text, and other content for the story. He should import the images and audio into Photo Story. In this stage, the storyteller can modify the number of images and image order, if necessary.

In the third stage, the storyteller creates, records, and finalizes the story. He should decide on the purpose and point of view of the story and write a script that will be used as narration. He can record the narration with a computer microphone and import the narration into Photo Story. Finally, the digital story is finalized by saving it as a Windows Media Video (.wmv) file.

In the fourth stage, the storyteller presents the story and receives feedback. He can show the story to his colleagues and gather feedback about how the story could be improved, expanded, and used in the classroom. If well received, he can teach colleagues how to create their own digital story.

Digital storytelling has the potential to facilitate teaching and learning in the classroom. Consequently, many of teachers intend to utilize the technology in classrooms at all levels of schools, from K-12 to higher education. Digital storytelling, when it is integrated into the classroom setting, can be a compelling teaching method to gain and hold students' attention. At the same time, it provides a creative and open-ended environment (Sadik, 2008). It is worth considering the application of digital storytelling in various subjects. When digital storytelling is integrated with

content areas and across the curriculum, it can be a way of expressing students' ideas and thoughts (Gakhar & Thompson, 2007; Lee, 2005; Shin & Park, 2008). This expression can promote learners' active participation and emphasize the active role of students rather than teachers. Therefore, it encourages student-centered learning. The following figure depicts how digital storytelling enhances the convergence of four student-centered learning strategies: student engagement, reflection for deep learning, project-based learning, and technology integration into the classroom (Barrett, 2006). These learning strategies can be facilitated through digital storytelling activities. All of the four strategies are attracting a great deal of attention from educators and teachers nowadays.

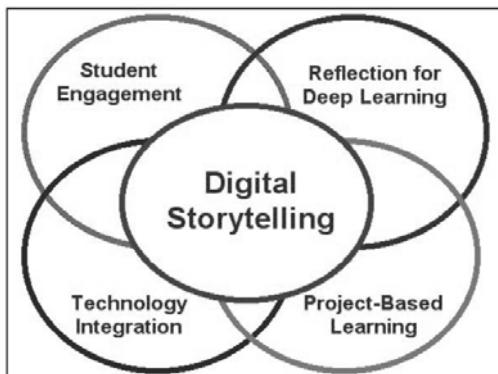


Figure 1. Convergence of student-centered learning strategies (Barrett, 2006)

Han (2007) contends that digital storytelling is a necessary element in producing edutainment learning contents. Edutainment is a core concept in creating learning contents in the 21st century. The word "edutainment" consists of "edu-" which is shortened from 'education', and "-tainment" which is shortened from 'entertainment'. Baek (2005) states that with the rapid development of computer technology and multimedia elements, edutainment can make learning much more attractive, potentially facilitating learning and eliciting students' active participation.

Students must master a wide variety of skills in order to accomplish the creation of a digital story, including narrative styling, written composition, visual literacy, media literacy, as well as an understanding of film conventions (Banaszewski, 2005). It is challenging for students to create their own digital stories and simultaneously acquire such skills. Although challenging, we can expect digital storytelling to foster such skills.

Research on digital storytelling continues. Sanchez (2009) asserts that the form of digital storytelling in Second Life is similar to general digital storytelling. The difference between them is the type of final products. The final products of general digital storytelling are video clips which are 2-5 minutes long. However, when it comes to Second Life, the artifacts are objects which are used to present the stories, giving texture to the objects.

In the research administered by Sanchez (2009), the students were asked to write a story, use tools in Second Life to make a digital story, and finally present their stories on the wall inside a condo in Second Life which was made by the researcher beforehand. The audience avatars walked through the condo, and the student who made the digital story gave a narration of his story. Sanchez asserts that digital storytelling is a way of self-expression. He points out that when students are making digital stories in Second Life, they can see each other's work. He proposes that this is a type of asynchronous social learning and modeling.

On May 17, 2008, the Center for Digital Storytelling and the New Media Consortium held a conference in Second Life about "Storytelling in Virtual Environments". Jen Friedberg, one of the speakers at the conference, said that storytelling in virtual worlds is beneficial. Accordingly, in virtual worlds, components of a story such as images and learning contents can be presented as realistic and vivid, almost similar to the real world. The users who see the digital story will have a strong sense of being there, so that will make it difficult to feel that they are using a computer. She suggests that it is advisable to properly utilize new powerful tools to present stories.

Kelleher and Pausch (2005) utilized a program called "Storytelling Alice" to inspire middle school girls' interest in programming. Storytelling Alice is a tool for creating 3D virtual world as a form of storytelling. Shin and Park (2008) conducted a study about digital storytelling in a virtual world, and verified the effect of digital storytelling type on learners' involvement and comprehension. Three scenarios of digital storytelling were performed: one where

the audience only listened, one where the audience listened and interacted, and finally one where the audience participated in constructing the environment of the story. The results showed that students were more engaged and enthusiastic in the third scenario. The study demonstrates that the realistic story environment of the 3D virtual world made a significant contribution to students' active engagement. Ibanez, Aylett, and Ruiz-Rodarte (2003) proposed storytelling in virtual environments from a virtual guide perspective. The guide navigates the virtual environment by storytelling.

Writing in Digital Storytelling

Although the ultimate purpose of digital storytelling is to tell the story to the audience, story writing is also important. A good script makes the digital story more effective and successful. Robin (2007) stresses story writing, describing how people usually spend more "digital" time searching for images and audio files rather than on the story itself. In contrast, he emphasizes that the process of selecting a meaningful topic and writing the story about the topic is the most important process in digital storytelling. When students pay more attention to the writing process, they will have ownership of their stories. Likewise, students will be motivated and engaged in the process of digital storytelling. In summary, although digital elements are important, writing should not be neglected, because it is an essential part of digital storytelling.

Gakhar and Thompson (2007) point out that digital storytelling provides students with opportunities to write creatively and organize thoughts in coherent ways, by designing and producing realistic artifacts. They also demonstrate that digital storytelling can improve students' writing skills, critical thinking skills, and media literacy.

Ballast, Stephens, and Radcliffe (2008) utilized digital storytelling as a means of improving students' writing. They examined the effects of digital storytelling on sixth grade students' writing and their attitudes toward writing. The experimental group was asked to craft a digital story with Photo Story. The control group was asked to convert a story which was written in advance into a word document, and then upload it to a class online magazine. At the end of the course, the researcher interviewed the students about the revisions they had made in the process of creating their digital stories. According to the interview data, the researchers found that the students in the experimental group were more engaged in digital storytelling and paid more attention in revising their writing compared to the control group who did not create digital story. The traditional writing group edited their writing merely by checking spelling and counting words. The digital storytelling group edited their stories more carefully by choosing appropriate words, adjusting the sequence of sentences, or removing sentences.

Warburton and Perez-Garcia (2009) reviewed multi user virtual environments, mainly Second Life. They examined the advantages of virtual environments and Second Life. First, virtual worlds can assist in visualizing and contextualizing objects we cannot see because of distance and monetary restrictions, or things which cannot be seen in real life situations. Second, there are a variety of interactions, including avatar-to-avatar interaction and avatar-to-object interaction. Third, the user is presented as an avatar, feeding the sense of really being in a virtual environment, unconscious of the mediation of the computer. Such feelings of immersion in virtual worlds enrich learners' experiences.

Digital storytelling gives expression to the ancient art of storytelling in a modern way. Multimedia components, such as images, sound, music, and others, are added to present the story. General digital storytelling is edited with software, such as Windows Movie Maker, Apple iMovie, and Adobe Premiere. However, the 3-D virtual world of Second Life maximizes the effects of multimedia, potentially employing components such as 3-D objects, images (snapshots), sounds, videos, and others. According to Sanchez (2009), in Second Life, digital storytelling follows the same premise as general digital storytelling. However, the key difference is that instead of creating a video, students create an experience or a space in which others can walk through their stories.

In spite of the educational potentials that virtual environments have, little research has been conducted about teaching writing in virtual environments, such as games and virtual worlds. In research conducted by Warren and Dondlinger (2008), a multi-user virtual environment named Anytown was used to support students' writing, on the basis of Problem-Based Learning. In Anytown, learners are required to take the role of cub reporters. A series of mysterious events were embedded in this MUVE including a burning building and strange lights emanating from the town's river. Learners should investigate these events in order to complete their writing tasks. The findings from this

research showed that learners acquired improvements in their writing performance in comparison to the traditional form of writing. Mayo (2004) developed a multi-player interactive game called Story world Builder, which was designed to motivate learners to write better. In this game, learners design and build a virtual place as the environment for their stories. They play active roles in the virtual environment as characters within the story. The program saves the transcript. The students write their stories using the transcript.

Research questions

This study aims to better understand the impact of digital storytelling on student learning, in particular when students actively engage in the storytelling process. The object is to find a meaningful difference in writing for digital storytelling on writing self-efficacy and flow between learners who use virtual reality learning environment and those who use Windows Movie Maker off-line.

Research Methods

Participants, settings and procedures

The target population for this study consisted of sixty-four undergraduate university students in South Korea. They were from two classes taking “Media Production for Teaching and Learning” at the Korean National University of Education. There were 32 students in each class. Their ages ranged from 20 to 22 years old. One class was selected to perform digital storytelling in the virtual world of Second Life, and the other class was selected to perform digital storytelling off-line utilizing the video editing software Windows Movie Maker.

In this study, participants were divided into two groups. One group performed digital storytelling off-line, utilizing Windows Movie Maker. Another group performed digital storytelling in Second Life. The off-line group followed the procedures of general digital storytelling suggested by the University of Houston. The activity for the Second Life group was designed by the author of this study. The procedures were as follows:

Step 1) Introduction to Second Life

Ask students to make an account of Second Life. Then, introduce the basic functions of Second Life, such as walking, flying, teleporting, communicating, object building. Also teach students how to search for freebie items, how to upload images, how to take snapshots, and how to buy items.

Step 2) Introduction to Digital Storytelling

Teach the definition and procedures (step 3 to step 7) of digital storytelling, and demonstrate an example of digital storytelling in Second Life.

Step 3) Story topic: Travelling through spaces and time

The topic of the story is travelling through spaces and time. They can create the story; it can be real or imaginary.

Step 4) Imagination of the story

The students do team work. A team consists of five to six students. Because there are 32 students in a class, four teams consist of five members, and the other two teams consist of six members. In this step, students travel through the places in Second Life that they like and capture photos of the places they plan to use in their story. They should get ideas from the places they visit in Second Life, and then gather ideas to compose a story.

Step 5) Writing the story

On the basis of step 4, students write down the story in a Word Processing file.

Step 6) Creating the environment of the story in Second Life

After writing the story, pick out the main sentences and add them into the images they captured. Then, build panels and upload the images to create the environment. Finally, arrange the panels in the sequential order, so that other Second Life residents can walk through the same path and see the story. Throughout this process, the learners should try to uncover some shortcomings in their writing, continuously going back to step 5 to modify their stories. Then they can create the environment again, according to the modified story. Repeat this process until the story is completed.

Step 7) Sharing stories

In this step, each team tells their story. One student should narrate. Other students should be the audience. Afterwards, they can give advice to each other.

Instruments

The primary tool used in this study was the virtual world called Second Life. Additionally, two kinds of tests were administered: first, the test of writing self-efficacy, and second, the test of flow state. Both of them included a pretest and a posttest. Each test was administered twice, before and after the experiment. Details about the tests are described below.

Second Life virtual learning environment for digital storytelling

Second Life is a 3D virtual world imagined and created by its residents. Since it was launched in 2003, Second Life has attracted attention from educators all over the world as an instructional technology platform (Park et al., 2008; Sanchez, 2009). Academic institutions from at least 14 countries are engaged in Second Life. The number of institutions is more than 150 (Foster, 2007). Recently, Second Life has been the most popular multi-user virtual environment used in education (Warburton & Perez-Garcia, 2009). Residents in Second life are presented as avatars. They can easily build, walk, fly, and teleport to anywhere they want to go in the virtual environment. There are also communication tools, like Instant Message, Local Chat, Note card, and Voice Chat. The functions integrated with its interface make it possible to build an environment in which digital storytelling can take place.

Test of Writing Self-Efficacy (Pre-post Test)

The writing self-efficacy questionnaire by Pajares and Valiante (2001) was used in this study. The questionnaire consists of 10 questions measured on a ten-point Likert scale. The total score of this test is 100 point. The reliability of this test is .88 (Cronbach's α). In this study, the questionnaire was modified and translated into Korean.

Table 1. Writhing self-efficacy test

| Content of Item |
|--|
| 1. I can write down my opinion without wandering away from the subject. 2. I can make the story that contains a subject have powerful impression. 3. I can construct simple sentences using appropriate grammar. 4. I can write a well organized and systemized article, which consists of introduction, main body, and conclusion. 5. I can complete a sentence with a proper conclusion. 6. I can appropriately use verbs, prepositions, adjectives, and so on. |

- | |
|--|
| 7. I can construct appropriate dialogue. |
| 8. I can use correct words to make a sentence |
| 9. I clearly know the place where should I place emphasis in a sentence. |
| 10. I can structuralize paragraphs, to verify the main idea of the topic sentence. |

Flow State Scale (Pre-post Test)

The flow scales are self-report instruments designed to assess the construct of flow, or optimal experience. The Flow State Scale was developed by Jackson and Eklund (2004). The scale was developed on the basis of Csikszentmihalyi (1990)'s flow theory. It contains of 36 items assessed on a five-point Likert scale. The reliability (Cronbach's α) of the test in this study is .95.

Research Design

A quasi-experimental analysis was carried out in two classes at Korea National University of Education, during 6 weeks in the second semester of the 2009 academic year. In this study, the activity of writing for digital storytelling was the independent variable. The dependent variables were writing self-efficacy and flow. There were two groups in this study. One group performed digital storytelling in Second Life, and another group created their digital stories off-line. Two-independent samples t-tests and multivariate analysis were used to compare the changes in writing self-efficacy between two groups, and also changes in flow score between two groups.

| | | | | | |
|----------------|----------------|----------------|----------------|----------------|----------------|
| G ₁ | O ₁ | O ₂ | X ₁ | O ₃ | O ₄ |
| G ₂ | O ₅ | O ₆ | X ₂ | O ₇ | O ₈ |

G₁ : Second Life Group / G₂ : Off-line Group

O₁, O₅ : Writing Self-Efficacy Pre-test/O₂, O₆ : Flow State Pre-test

X₁ : Writing for Digital Storytelling in Second Life / X₂ : Writing for Digital Storytelling Off-line

O₃, O₇ : Writing Self-Efficacy Post-test / O₄, O₈ : Flow State Post-test

Results

The main purpose of this study was to examine the effects of writing for digital storytelling on writing self-efficacy and on flow in the virtual world of Second Life. In this study, one of the groups created their digital stories in Second Life, and the other group created their digital stories off-line with Windows Movie Maker. Before and after the activity of digital storytelling, students' writing self-efficacy and flow were measured using the tests mentioned above in the research methodology section. Two-independent sample *t*-tests and multivariate analysis were performed using the SPSS 17.0 statistics program. The changes in scores of writing self-efficacy between the two groups were compared, and the changes in scores of flow between the two groups were also compared. This test was found to be statistically significant. The following table summarizes the results of the two-independent samples *t*-tests and multivariate analysis.

Table 2. Results of two-independent samples t-tests

| | Group | N | M | SD | t | p |
|-----------------------|-------------------|----|------|------|--------|----------|
| Writing Self-Efficacy | Second Life Group | 32 | 6.63 | 7.76 | 2.307 | 0.024* |
| | Off-line Group | 32 | 2.50 | 6.49 | | |
| Flow | Second Life Group | 32 | 3.67 | .16 | 13.641 | 0.001*** |
| | Off-line Group | 32 | 3.07 | .19 | | |

* Correlation is significant at the 0.05 level. *** Correlation is significant at the 0.001 level.

As indicated in Table 2, the difference in the mean change score of writing self-efficacy between the two groups was significant, $t(62)=2.307$, $p<.05$. The mean change score of writing self-efficacy in the Second Life group ($M=6.63$, $SD=7.76$) was significantly higher than that of the off-line group ($M=2.50$, $SD=6.49$). The alpha level was .05. The total mean score of writing self-efficacy is 10 point. The mean change score has been improved 4.13 point. Namely, the activity of writing for digital storytelling was effective in improving students' writing self-efficacy.

The difference in the mean change scores of flow between the two groups was significant, $t(62)=13.641$, $p<.001$. The mean change scores of flow in the Second Life group ($M=3.67$, $SD=.16$) was higher than that of the off-line group ($M=3.07$, $SD=.19$). The alpha level was .001. The total mean score of flow is 5 points. The mean change score has been improved .60 point. It indicates that the activity of writing for digital storytelling in Second Life is effective in enhancing students' flow level. The experiment results demonstrate that digital storytelling in Second Life is more effective than digital storytelling off-line.

The flow scales assess 9 dimensions of flow. The scale consists of 9 sub-categories containing a total of 36 items. *From flow scale characteristic* there is a close correlation among sub-category variables of flow. We used multivariate analysis to verify the variations in flow level among sub-category variables. Results of the post-flow multivariate analysis of the relationship between the groups are displayed in Table 3. According to the result, there was significant difference between virtual reality learning group looks like Second Life and off-line with movie maker for digital storytelling ($\lambda=.003$, $p=.001$). In other words, the group that performed the digital storytelling activity in Second Life experienced a higher level of immersion than the group that did the digital storytelling activity with Windows Movie Maker.

*Table 3. Results of multivariate analysis for flow (Wilks' $\lambda=.003^{***}$, $p<.001$)*

| Variance Sources | | <i>M</i> | <i>SD</i> | <i>F</i> | <i>p</i> | η^2 |
|---------------------------------|-------------------|----------|-----------|------------|----------|----------|
| Challenge-Skill Balance | Second Life(n=32) | 3.69 | .25 | 117.474*** | .001 | .655 |
| | Movie Maker(n=32) | 2.94 | .30 | | | |
| Merging of Action and Awareness | Second Life(n=32) | 2.53 | .12 | 72.145*** | .001 | .538 |
| | Movie Maker(n=32) | 2.03 | .31 | | | |
| Clear Goals | Second Life(n=32) | 3.86 | .16 | 46.721*** | .001 | .430 |
| | Movie Maker(n=32) | 3.50 | .25 | | | |
| Unambiguous Feedback | Second Life(n=32) | 3.84 | .20 | 119.370*** | .001 | .658 |
| | Movie Maker(n=32) | 3.23 | .25 | | | |
| Total Concentration | Second Life(n=32) | 3.45 | .27 | 4.740* | .033 | .071 |
| | Movie Maker(n=32) | 3.30 | .30 | | | |
| Sense of Control | Second Life(n=32) | 3.36 | .16 | 18.036*** | .001 | .225 |
| | Movie Maker(n=32) | 3.48 | .06 | | | |
| Loss of Self-Consciousness | Second Life(n=32) | 3.98 | 1.03 | 117.952*** | .001 | .655 |
| | Movie Maker(n=32) | 3.57 | .38 | | | |
| Time Transformation | Second Life(n=32) | 4.33 | .38 | 62.984*** | .001 | .504 |
| | Movie Maker(n=32) | 3.58 | .38 | | | |
| Autotelic Experience | Second Life(n=32) | 3.94 | .25 | 10.145** | .002 | .141 |
| | Movie Maker(n=32) | 3.56 | .62 | | | |

* Correlation is significant at the 0.05 level. ** Correlation is significant at the 0.01 level. *** Correlation is significant at the 0.001 level.

From the results of the Multivariate Analysis of Variance (MANOVA), among the nine categories of flow (challenge-skill balance, merging of action and awareness, clear goals, unambiguous feedback, total concentration, sense of control, loss of self-consciousness, time transformation, autotelic experience), a significant difference existed among digital storytelling groups (Wilks' $\lambda = .003$, $p<0.001$). Wilks' λ represents the part of within-group variation against total variation, where a Wilks' λ close to 0 means that differences between groups are major. According to η^2 , which represents the explanatory power of the independent variable, 66% of total change score can

be explained as the contribution of challenge-skill balance, 54% as the contribution of merging of action and awareness, 43% as the contribution of clear goals, 66% as the contribution of unambiguous feedback, 7% as the contribution of total concentration, 23% as the contribution of sense of control, 66% as the contribution of loss of self-consciousness, 50% as the contribution of time transformation, 14% as the contribution of autotelic experience. The main factors contributing to digital storytelling in virtual worlds are challenge-skill balance, unambiguous feedback, and loss of self-consciousness.

The virtual environment like Second Life can provide players balanced challenge and skill. This makes players motivated and engaged in the activity of digital storytelling. In virtual worlds, students can see their stories intuitively, in the whole process of creating digital stories. Thus, they can find problems in their stories, and make revisions more easily. This can be considered as a kind of intrinsic feedback. Learners are represented as avatars, and who participated in the digital storytelling activity is also the avatar. Avatar is the agents of students. Therefore, when students are totally involved in the activity in the virtual world, they can't even be conscious of themselves. All three elements are an important factor that leads to flow, in the activity of digital storytelling in virtual worlds.

Conclusion and Suggestions

Change increasingly defines the nature of storytelling in an information age. Storytelling is rapidly and continuously changing as new technologies for information and communication emerge and users craft new methods for employing these technologies. Moreover, these new technologies for information and communication permit the exchange of even newer technologies and visions for their use. This speeds up the already rapid pace of change in the forms and functions in virtual worlds, increasing the complexity of the challenges we face as we consider how to prepare students for their virtual reality environment.

Meaningful technological integration is defined as curricula utilizing authentic tasks that intentionally and actively help learners to construct their own meanings from thinking about experiences (Jonassen et al., 1999). Digital stories are a form of communication and as they are created, students apply critical thinking skills while selecting the appropriate media to convey the story's message to the audience. It provides the student with a learning environment to apply communication skills, work collaboratively, and think critically while addressing content and technology standards.

The purpose of this study was to examine the effects of writing for digital storytelling on writing self-efficacy and on *flow* in a virtual learning environment such as Second Life. Through statistical analysis, we can conclude that digital storytelling in Second Life is more effective than digital storytelling off-line. Writing self-efficacy and flow can be improved through actively engaging in writing for digital storytelling in a virtual-reality learning environment similar to Second Life.

The findings of this study have implications for teachers of writing. Digital storytelling can be used in classroom settings as a means of teaching writing. In the virtual world, students can visit virtual places using the teleporting function, without any constraint of time and space. It can stimulate their imaginations and encourage more creative writing. Then they can visualize the stories they imagine, by creating objects and changing the forms of those objects, thereby realistically enacting their stories. Through the visualization, they can see the structure of the whole story clearly. It will help them to identify and correct mistakes they have made, to construct a solid logical structure, and to rearrange their thinking in a logical way.

As virtual reality technology has advanced, it has been widely touted as a major advance that can offer meaningful support for educational purposes. There are several ways in which VR technology is expected to facilitate learning. One of its unique capabilities is the ability to provide students with the opportunity to learn in open-ended learning environments and to interact in ways that distance, time, or safety factors make unavailable. The type of activities supported by this capability allows for the prospect that students will be better able to master, retain, and share new knowledge as they actively engage in constructing knowledge in virtual learning environments. Jakes and Brennan (2005) propose that, in the future, digital storytelling can be a potent learning experience that encompasses much of what society hopes that students will know and be able to perform in the future classroom.

References

- Armstrong, S. (2003). The power of storytelling in education. In Armstrong, S. (Ed.), *Snapshots! Educational insights from the Thornburg Center* (pp. 11-20). The Thornburg Center: Lake Barrington, Illinois.
- Baek, Y. K. (2005). *Understanding and Utilization of Edutainment*. Seoul: Jung-il.
- Ballast, K., Stephens, L., & Radcliffe, R. (2008). The Effects of Digital Storytelling on Sixth Grade Students' Writing and Their Attitudes about Writing. In K. McFerrin et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2008* (pp. 875-879). Chesapeake, VA: AACE.
- Banaszewski, T. (2005). *Digital storytelling, supporting digital literacy in grades 4-12*. A thesis presented to the academic Faculty, Georgia Institute of Technology. Retrieved August 3, 2009, from <http://hdl.handle.net/1853/6966>.
- Barrett, H. (2006). Convergence of student-centered learning strategies. *Technology and Teacher Education Annual*, 1, 647-654.
- Chung, S. K. (2006). Digital Storytelling in Integrated Arts Education. *The International Journal of Arts Education*, 4(1), 33-50.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper & Row.
- Digital Storytelling Association (2002). *The center for digital storytelling*. Retrieved 18 February, 2009, from <http://www.dsaweb.org>.
- Dickey, M.D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development*, 53(2), 67-83.
- Foster, A. L. (2007). Professor avatar. *Chronicle of Higher Education*, 54(4), 24-26.
- Gakhar, S., & Thompson, A. (2007). Digital Storytelling: Engaging, Communicating, and Collaborating. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2007* (pp. 607-612). Chesapeake, VA: AACE.
- Gregori-Signes, C. (2008). Integrating the Old and the New: Digital Storytelling in the EFL Language Classroom. *GRETA Journal*, 16(1), 43-49.
- Han, H. W. (2007). Theory and Practice of Digital Storytelling. *The society of Korean Language and Culture*, 32, 27-43.
- Ibanez, J., Aylett, R., & Ruiz-Rodarte, R. (2003). Storytelling in virtual environments from a virtual guide perspective. *Virtual Reality*, 7(1), 30-42.
- Jackson, S. A., & Eklund, R. C. (2004). *The flow scales manual*. Morgantown, WV: Fitness Information Technology.
- Jakes, D. S., & Brennan, J. (2005). Capturing stories, capturing lives: An introduction to digital storytelling. Retrieved February 8, 2010, from http://www.jakesonline.org/dstory_ice.pdf.
- Jonassen, D., Peck, K., & Wilson, B. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Prentice Hall.
- Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys (CSUR)*, 37(2), 83-137.
- Lee, I. H. (2005). *Korean Style Digital Storytelling*. Seoul: Sallim Knowledge.
- Kang, S. H., Ko, U., Bae, J. Y., Lee, I. H., Lee, J. Y., Jun, K. R., Jun, B. K., & Han, H. W. (2003). *Digital Storytelling*. Seoul: Golden Bough.
- Mayo, M. (2004). A multi-player educational game for story writing. *Paper presented at the Australia Workshop on Interactive Entertainment*, 13 February, Sydney.
- McDrury, J., & Alterio, M. (2003). *Learning through storytelling in higher education*. Sterling, VA: Kogan Page.
- Mello, R. (2001). The power of storytelling: How oral narrative influences children's relationships in classrooms. *International Journal of Education & the Arts*, 2(1). Retrieved March 19, 2009, from <http://ijea.org>.
- Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology*, 26, 366-381.
- Park, H. S., & Baek, Y. K. (2009). Empirical evidence and practical cases for using virtual world in educational context. In H. Yang & S. Yuen (Eds.), *Collective intelligence and e-learning2.0: Implications of web-based communities and networking* (pp. 227-246). Hershey, PA: IGI Global.

- Park, H. S., & Jung, J. Y., & Chris, C. (2008). Effect of the activities in the Second Life as tool for mutual understanding of gender role. In K. McFerin, R. Weber, R. Carlsen & D. A. Willis (Eds.), *Proceedings of the Society for Information Technology and Teacher Education*, USA, 19, 1778-1783.
- Park, E. J., & Seo, J. H. (2009). Applying digital storytelling technique to website navigation for improving emotional user experience. *Proceeding of the International Association of Societies of Design Research 2009*, Seoul, Korea, 4125-4128.
- Paul, N., & Fiebich, C. (2002). The elements of digital storytelling. Retrieved August12, 2009, from <http://www.inms.umn.edu/elements/index.php>.
- Robin, B. (2007). The Convergence of Digital Storytelling and Popular Culture in Graduate Education. In R. Carlsen et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2007* (pp. 643-650). Chesapeake, VA: AACE.
- Robin, B. (2008). Digital Storytelling: A Powerful Technology Tool for the 21st Century Classroom. *Theory Into Practice*, 47, 220-228.
- Sadik, A. (2008). Digital storytelling: a meaningful technology-integrated approach for engaged student learning. *Educational Technology Research and Development*, 56(4), 487-506.
- Sanchez, J. (2009). Pedagogical Applications of Second Life. *Library Technology Reports*, 45(2), 21-28.
- Shin, B. J., & Park, H. S. (2008). The effect of digital storytelling type on the learner's fun and comprehension in virtual reality. *Journal of the Korean Association of Information Education*, 12(4), 417-425.
- University of Houston (2009). *The Educational Uses of Digital Storytelling*. Retrieved June 14, 2009, from <http://digitalstorytelling.coe.uh.edu>.
- Warburton, S., & Perez-Garcia, M. (2009). 3D design and collaboration in massively multi-user virtual environments. In D. Russel (Ed.), *Cases on collaboration in virtual learning environments: processes and interactions* (pp.27-41). Hershey, PA: IGI Global.
- Warren, S. J., & Dondlinger, M. J. (2008). A MUVE Towards PBL Writing: Effects of a Digital Learning Environment Designed to Improve Elementary Student Writing. *Journal of Research on Technology in Education*, 41(1), 113-140.

Exploring Gender Differences in SMS-Based Mobile Library Search System Adoption

Tiong-Thye Goh

School of Information Management, Victoria University of Wellington, New Zealand // tiong.goh@vuw.ac.nz

ABSTRACT

This paper investigates differences in how male and female students perceived a short message service (SMS) library catalog search service when adopting it. Based on a sample of 90 students, the results suggest that there are significant differences in perceived usefulness and intention to use but no significant differences in self-efficacy and perceived ease of use between genders. The findings reveal that SMS efficiency has a significant influence on self-efficacy for males but not for females. Online Public Access Catalog (OPAC) experience has a stronger but less significant influence on self-efficacy for females but very little influence for males. Perceived usefulness is still the driving force behind intention to use for both genders. The findings suggest that factors influencing self-efficacy and its impact on Technology Acceptance Model (TAM) constructs are different for each gender. The implications of these findings are discussed for both researchers and practitioners.

Keywords

Short message service, Technology adoption, Mobile search, Self-efficacy, Gender differences

Introduction

According to the Australian and New Zealand Institute for Information Literacy (ANZIIL) framework (Bundy, 2004), one of the core competencies for the information literate person is the ability to find needed information effectively and efficiently. Competency in information literacy through the learning of fundamental library skills such as information seeking is critical for academic success and survival (Macpherson, 2004). In addition, to be a lifelong learner one needs to demonstrate and acquire the ability to retrieve information using a variety of media (Candy et al., 1994). Hence individual learners, both male and female, who can become proficient with new search technology will be well prepared both in school and the wider society.

Previous research has identified gender to be a strong predictor of attitudes and behavior in internet information seeking and a demographic factor that strongly influences information behavior (Weiser, 2000; Roy & Chi, 2003). These behaviors include verbal and learning skills, spatial orientation and interest in computers (Hupfer & Detlor, 2006). However limited research has been done on gender differences in mobile library search systems. Understanding gender differences allows us to develop better strategies and systems to assist individual learners to acquire needed information literacy skills and contribute to better learning experiences and academic success (Macpherson, 2004).

Research has identified gender differences in the use of technology and especially in information communication technology (ICT) (Selwyn, 2007). For example research has shown that females are normally less confident than males in handling technology such as computer software. The research conducted by Madigan et al. (2007) to compare students' skills in four major technology domain areas indicated that first year female students showed less confidence in using computer technology and they did not perceive themselves as competent technology users. Being less confident will have implications for how individuals perceive self-efficacy toward ICT. The research by Hartzel (2003) identified that female students tend to have a lower self-efficacy than male students while using software packages. In software training, the genders also differ in training satisfaction. Bradley and Lee (2007) reported that females had significantly lower levels of satisfaction with enterprise resource planning (ERP) training and they wanted to receive more training than males. In the context of mobile service adoption, Sohn and Lee (2007) discovered females are more inclined to adopt text messages than males, perhaps due to self-presentation and willingness (Sohn & Lee, 2007). Evidently the different behaviors can significantly affect the ways in which technology enhanced learning, organization training and services should be delivered in order to achieve better acceptance, satisfaction and engagement from learners.

Using an extension of the Technology Acceptance Model (TAM), Gefen and Straub (1997) in their study on the perception and use of email indicate that women and men differ in their perceptions of social presence, usefulness, and ease of use but not in the use of email. Their investigations were based on gender having a direct impact on the four dominant antecedents but not on how each gender moderates the impact on the use of email. Venkatesh and

Morris (2000) extended the Technology Acceptance Model to study the moderating effect of gender on software use. Their findings suggest that men treat perceived usefulness as an important factor in deciding to use software while women treat perceived ease of use as the dominant factor in deciding to use software. In a closely related study of a mobile chat service, Nysveen et al. (2005) investigated the moderating effect of gender by including the Theory of Reasoned Action in the Technology Acceptance Model. Their findings suggest that females treat social norms and intrinsic motives as important determinants in using mobile chat services but ease of use has little impact on their intention to use. Amin (2007) adopted a model similar to that of Nysveen et al. (2005) in a study of a SMS mobile banking service. The study revealed that perceived expressiveness, perceived ease of use and social norms are the important determinants for intention to use SMS banking among female students. Conversely, perceived usefulness and perceived system quality are the important determinants for using SMS banking among male students.

These studies clearly demonstrate that research on gender and mobile technology adoption theory is an important aspect of information systems and learning research but has not yet been fully explored. In particular research is limited in the context of learning support and mobile information search. Therefore we propose to investigate gender differences with regard to a SMS based library catalog search system using a modified Technology Acceptance Model. We anticipate that it will lead to further research on gender issues in other forms of mobile search technologies that support learning and information literacy skills. By understanding the gender effects, our aim is to develop appropriate strategies that enhance the adoption and diffusion of mobile search services as an essential channel to support learning activities. SMS technology can be an effective platform for health education, mobile and blended learning, and for effective learning support in developing countries (Nokia, 2010; UN Foundation, 2009) and its potential should not be overlooked.

SMS-Based Services within the Library Sector

Academic libraries need to offer innovative services in order to provide quality service to students and adult learners (Lukasiewicz, 2007). For example, various enquiry channels such as telephone, email, and instant messages have been integrated into reference services (Hill et al., 2007) to provide better support to learners. With the immense popularity and adoption of text messaging services, SMS is now being used in many reference services to provide an additional channel of service. For instance, the Southbank Institute of Technology in Australia adopted a SMS reference service (Herman, 2007). Using MessageNet (2008) students can send SMS queries to the service section. The SMS is routed to a common email mailbox and service staff can reply to the user using the email to SMS facility. Two advantages of using SMS are that it could alleviate any spoken language difficulties of international students, and that it is part of youth culture. Another library pioneering the use of SMS technology for its services is Curtin University Library. Curtin adopted a “SMS a Query” system which allows users to send a query to the reference section via SMS (Giles & Grey-Smith, 2005). The message is also routed to a mailbox and staff can reply to the query directly from the email to SMS facility. Overall, users have responded well to the SMS query for reference service. A slightly different solution is the SMS to Web-based queries system. Adopted by the Helsinki University of Technology Library, the “library in your pocket” system utilized Liblet middleware that translates SMS messages into a Web-based queries system and vice versa (Pasanen & Muhonen, 2002). Unlike other SMS services, the system is a very comprehensive mobile service. It can send notices, reminders, reservations, and general queries, check the availability of individual items and allow renewal of borrowed items, and paying of fees. The service however required pre-registration. During its pilot phase it had more than 600 registered mobile users. Feedback from users was also very positive. A recent development is the SMS based “My Info Quest” (MIQ) collaborative distributed reference service which was reported to attract thousands of patrons (Jensen, 2010). While these examples demonstrate that SMS technologies have been gradually introduced in academic libraries they are definitely not widespread. A recent survey by Parker (2007) found only 50 English language libraries in the world adopting some form of SMS services. There is a need to further understand the adoption issues of SMS technologies in the library sector. This paper provides a timely insight into more specific gender differences in adopting a proposed SMS based library search and reservation system.

Mobile Catalog Search

Traditionally, library catalogs were searched through card catalogs, and more recently through Online Public Access Catalogs (OPAC). Searching for information via mobile devices is becoming increasing popular. Most recently

OPAC has been transformed into AirPAC (Auckland Council, 2010) for the use of mobile devices. Other Internet search engine providers such as Google and Yahoo provide general mobile search or specific mobile search using mobile browsers. Specific mobile search services such as location search, directory discovery search, and voice activated search, hybrid search and photo mobile search are gaining momentum (Fox, 2008). These kinds of search interface are all very similar to the desktop search interface. However due to small screen size, optimization and content adaptation are necessary to achieve better user experiences. Some of these mobile search services looked at the issue of mobility and context awareness (Mihalic et al., 2005; Gorlenko & Merrick, 2003). Other advanced services intend to communicate with, for instance, other devices, objects of interest or friends in the user's environment.

SMS text messaging has been widely adopted by the younger generation (Turel et al., 2007). The potential of SMS-based technology to enhance library services and engage with a younger generation of library users is evident. The system is also useful in developing countries where access to personal computers and the Internet is limited. The current study intends to investigate how each gender moderates the relationships and characteristics of SMS efficiency (SMSEFF), OPAC experience (OPAC), perceived self-efficacy (SE), perceived usefulness (PU), perceived ease of use (PEOU) and intention to use (IU) in relation to a proposed SMS-based library catalog system; we shall explain the details in a later section. Data will be collected via a survey of a purposeful and convenient sub sample. An early prototype of the SMS-based library catalog search system was developed for this purpose.

The SMS-Based Library Catalogue Search System Description

The current prototype system which was developed by the author is shown in Figure 1. The system consists of the user's mobile phone, a SMS center, an interface mobile phone, a SMS interface, a software agent, and OPAC search engines. The SMS center is operated by the telecom operator. When a user inputs a SMS search keyword to his mobile phone, the message is routed through the carrier SMS center (SMSC). The SMSC identifies the interface phone number and passes the message to the interface mobile phone connected to the SMS interface. Upon receiving the SMS message which includes the input (search) keywords, the SMS interface software parses the message from the interface mobile phone and activates the software agent to trigger the OPAC search. The software agent can activate several different search engines that are available. Once the search has been completed, the result is returned to the sender in the form of either a SMS message via the interface mobile phone or an email. Alternatively, if the Internet server is connected to the SMSC via TCP/IP, then upon receiving the search keywords from a user, the SMSC passes this information to the Internet server and subsequently activates the software agent to perform the search. The result is returned to the SMSC to be forwarded to the sender's mobile unit or sent as an email. In the long run, a direct connection between the SMSC and an Internet server is preferred. However, this would require the purchase of a SMSC account service with a carrier or a third party SMS service provider.

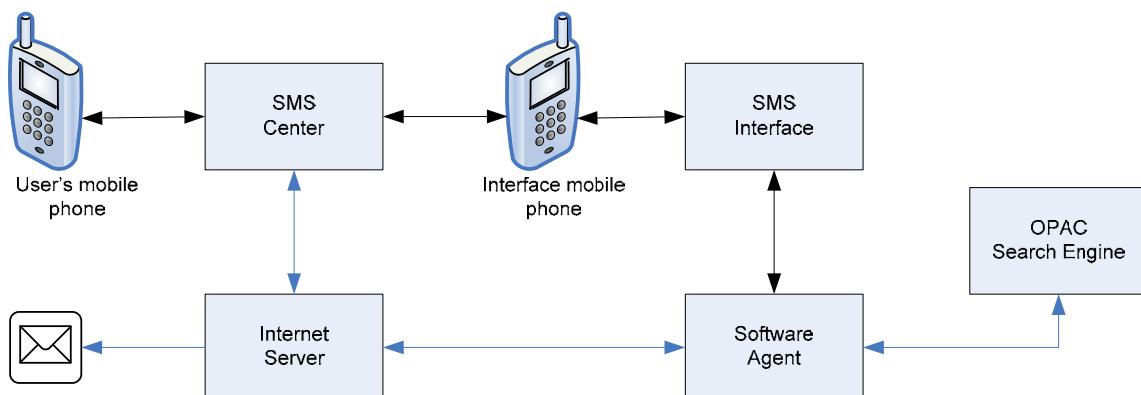


Figure 1. SMS search system

A series of screen shots of the prototype system is depicted in Figure 2 to show what will be displayed to the users via a mobile phone (or other type of mobile device).

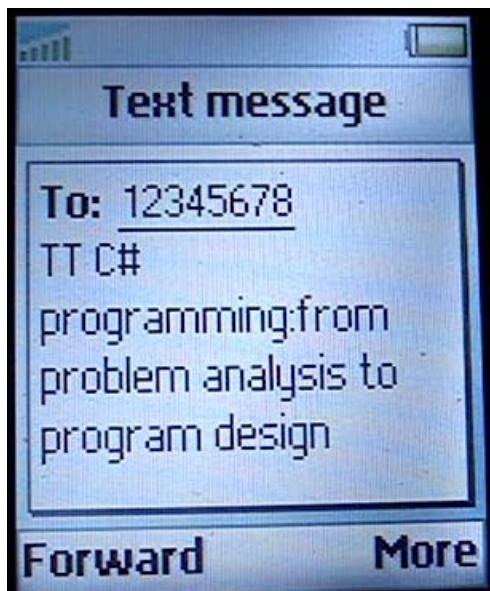


Figure 2(a) Title search

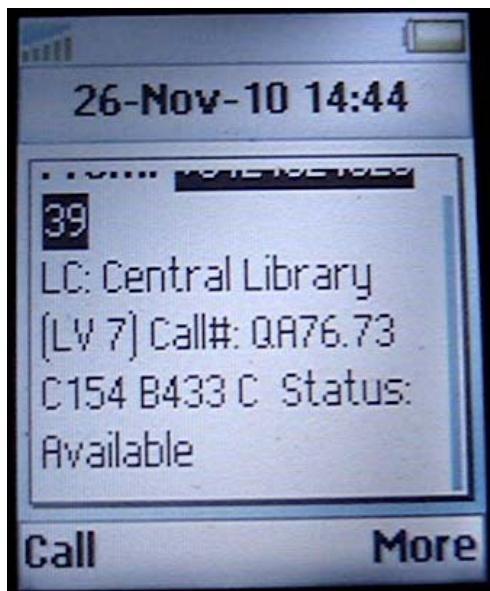


Figure 2(b) Display matched item



Figure 2(c) User's request a Hold



Figure 2(d) Confirmation of Hold

Figure 2. SMS search screen shots

Theoretical Development

The Technology Acceptance Model has been widely adopted as the basic framework for examining the adoption of various technologies (Lee et al., 2007; Park & Chen, 2007; Turel et al., 2007; Pedersen, 2006; Mallat et al., 2006; Lai, 2004; Dennis et al., 2003; Venkatesh et al., 2002; Kishore et al., 2001). Figure 3 shows the research model which was modified from the original TAM developed by Davis (1989). SE is an important factor in the use of electronic library systems (Ramayah & Aafaqi, 2004; Waldman, 2003); therefore this paper adds the SE into the research model. Moreover, according to Gefen and Straub (1997), gender difference also plays an important role in technology adoption; thus this paper adds gender as a moderator variable. The model hypothesizes that gender has a moderating effect on the relationships between OPAC experience and SE, SMSEFF and SE, SE and PEOU, SE and PU, SE and IU, PEOU and PU, PEOU and IU, and PU and IU.

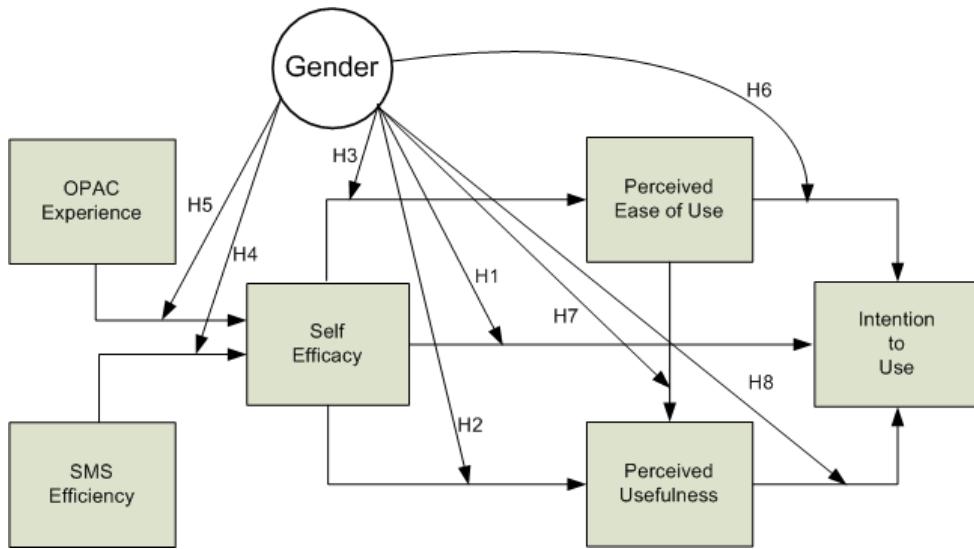


Figure 3. Hypothesis and model

Perceived Self-Efficacy

Wood and Bandura (1989, p. 408) provide a definition of perceived self-efficacy as “beliefs in one’s capabilities to mobilize the motivation, cognitive resources and courses of action needed to meet a given situation’s demands.” Self-efficacy is an important construct supported by social cognitive theory. An early study by Igbaria and Iivari (1995), who examined the influence of self-efficacy on computer usage, showed that SE had an indirect influence on usage through PEOU and PU. A study by Ramayah and Aafaqi (2004) on the impact of SE and cognitive beliefs toward technology adoption in e-library usage among students in a Malaysian university concurred that SE has a significant direct impact on PU and PEOU in e-library usage but no direct impact on usage. Venkatesh and Morris (2000) argued that women typically display lower SE, lower computer aptitude and higher computer anxiety than men. A higher level of anxiety can lead to a lower level of SE and subsequently influence outcome expectations (Compeau & Higgins, 1995). In the context of TAM, Chau (2001) argued that PU reflects a person’s beliefs or expectations about outcomes. Hence computer SE may be an important factor affecting PU. Therefore, a lower level of SE is likely to lead to lower levels of PEOU, PU and IU. Based on these arguments, we propose the following hypotheses:

- H1: SE will influence IU more strongly for females than males.
- H2: SE will influence PU more strongly for females than males.
- H3: SE will influence PEOU more strongly for females than males.

Previous Experience

According to Bandura (1994) one way that individuals form their self-efficacy beliefs is by constructing information from their previous experience. Previous experience has been found to be associated with SE. Task complexity can influence SE as indicated in a study by Chang (2005). In a study on software adoption Hartzel (2003) showed that women’s SE increased more than men’s after a more complex task-specific experience. In this study, the use of an OPAC can be argued to be more complex than sending SMS. The OPAC is a more complex system because of its various searching interfaces. Moreover SMS is widely used by both male and female students. Therefore one can argue that OPAC experience will have more salient impact on SE for females than males and skills in sending SMS will have less salient impact for females than males. Hence we hypothesize that:

- H4: SMSEFF will influence SE more strongly for males than females.
- H5: OPAC experience will influence SE more strongly for females than males.

Perceived Ease of Use

Davis (1989, p. 320) refers to perceived ease of use as “the degree to which a person believes that using a particular system would be free of effort.” SMS is used widely by both male and female university students and can be regarded as a simple tool. It is likely that both genders will assess the ease of use of the system equally. Venkatesh and Morris (2000) suggested that females have lower SE, lower computer aptitude and higher computer anxiety compared to men and proposed that the influence of perceived ease of use on the intention to use IT is stronger for females than for males. On the other hand, Ong and Lai’s (2006) findings suggested that PEOU is more salient in determining PU for females than males. Based on these arguments, we hypothesize that:

H6: PEOU will influence IU more strongly for females than males.

H7: PEOU will influence PU more strongly for females than males.

Perceived Usefulness

According to Venkatesh and Morris (2000) males tend to exhibit stronger and more sensitive attitudes toward task-oriented and instrumental applications of IT than females. Mobile search is a task-oriented activity and there is reason to believe that the influence of PU on IU regarding the search system is stronger for males than for females. We hypothesize that:

H8: PU will influence IU more strongly for males than females.

Methodology

Subjects

The target respondents of this study were university students who had experience in using the university’s OPAC and send SMS regularly. Table 1 summarizes the background information of the participants. A quantitative approach was taken and a survey questionnaire was the data collection instrument for this study. Prior to data collection, a series of screenshots similar to Figure 2 were displayed which explained the functions and features of the system. All returned questionnaires were manually checked to ensure there were no missing or ambiguous answers. A total of 112 questionnaires were received out of the 130 distributed, of which 9 were incomplete and were therefore excluded from the final analysis. A sample of 90, representing 87.3% of the population, was analyzed for the 16-25 year age group which represents the younger generation. Given the small sample size, structure equation modeling (SEM) was not advisable (Hair et al., 1998; Gefen et al., 2000). A sample of at least 30 is required for linear regression with appropriate power (Gefen et al., 2000; Cohen, 1992) while a sample size of 150-200 is recommended for SEM (Gefen et al., 2000). Hence linear regression is adopted for the analysis.

Measures

The questionnaire consisted of four parts. The first part comprised demographic questions. The second part focused on the cognitive belief constructs PU and PEOU. Under PU, six statements were presented and respondents were asked to rate them from 1 (very unlikely) to 7 (very likely) to investigate whether they perceived the proposed system to be useful. Another six statements were presented for PEOU of the proposed system and again, respondents were asked to rate them from 1 (very unlikely) to 7 (very likely). These statements were adapted from Davis (1989) and were to serve as an initial stimulus.

The third part of the questionnaire investigated the perceived SE of the respondents toward the proposed system. This part of the questionnaire was based on Compeau and Higgins’ (1995) work on computer SE that examined one’s perceived ability to use an unfamiliar piece of software. Ten statements/conditions were included in this part of the questionnaire to solicit respondents’ perception of whether they could use the proposed system under a variety of conditions. For each of the conditions, respondents were asked to indicate whether they thought they would be able to complete their task using the proposed system, and for each of the conditions for which they answered ‘yes’, they

were asked to rate their confidence in their original judgment by selecting a number from 1 (not at all confident) to 7 (totally confident).

Finally, the fourth part of the questionnaire investigated the participants' IU of the proposed system.

Table 1. Participants' background information

| <i>Variables</i> | <i>Frequency</i> | <i>Male</i> | <i>Female</i> | <i>%</i> |
|---------------------------|------------------|-------------|---------------|----------|
| Gender | N=90 | | | |
| Male | 55 | | | 61.1 |
| Female | 35 | | | 38.9 |
| Age | N=90 | | | |
| 16-20 | 67 | 42 | 25 | 74.4 |
| 21-25 | 23 | 13 | 10 | 25.6 |
| Year in university | N=90 | | | |
| 1 st | 58 | 36 | 22 | 64.4 |
| 2 nd | 23 | 12 | 11 | 25.6 |
| 3 rd | 5 | 4 | 1 | 5.6 |
| 4 th | 3 | 2 | 1 | 3.3 |
| 5 th | 1 | 1 | 0 | 1.1 |
| OPAC experience | N=89* | | | |
| >2 years | 21 | 12 | 9 | 23.3 |
| 1-2 years | 23 | 12 | 11 | 25.6 |
| <1 year | 45 | 30 | 15 | 50.0 |
| Mobile phone user | N=89* | | | |
| Never | 1 | 0 | 1 | 1.1 |
| 1-2 years | 2 | 2 | 0 | 2.2 |
| 2-4 years | 23 | 17 | 6 | 25.6 |
| 4-6 years | 32 | 17 | 15 | 35.6 |
| >6 years | 31 | 19 | 12 | 34.4 |
| SMS usage | N=90 | | | |
| Never | 3 | 2 | 1 | 3.3 |
| <5 msg / day | 6 | 5 | 1 | 6.7 |
| 5-10 msg / day | 29 | 17 | 12 | 32.2 |
| >10 msg / day | 52 | 31 | 21 | 57.8 |
| SMS efficiency | N=89* | | | |
| Novice | 2 | 1 | 1 | 2.2 |
| > Novice | 1 | 1 | 0 | 1.1 |
| < Intermediate | 4 | 2 | 2 | 4.4 |
| Intermediate | 13 | 5 | 8 | 14.4 |
| > Intermediate | 14 | 9 | 5 | 15.6 |
| < Proficient | 23 | 13 | 10 | 25.6 |
| Proficient | 32 | 23 | 9 | 35.6 |

* 1 missing entry

Data Analysis and Results

Analysis of Measurement Validity

The samples consisted of more male (61.1%) than female (38.9%) participants. Most (74.4%) of the respondents were from the 16-20 year age group and 25.6% were aged 20-25. Nearly all (90%) were first and second year students. Except for one participant, all the other participants had used a mobile phone and 96.7% had used SMS. A similar majority (91.2%) of the participants believed their texting skills were at an intermediate or higher level.

Internal reliability was validated through Cronbach's alpha. Cronbach's coefficient alpha in Table 2 shows that the internal reliabilities of the constructs were all greater than 0.8; a value of 0.7 or higher is normally considered acceptable (Hair et al., 1998). Both male and female samples were equally strong.

Table 2. Cronbach's alpha

| Factors | Items | Cronbach's alpha | | |
|------------------------------|-------|------------------|-------|--------|
| | | All | Male | Female |
| Perceived Ease of Use (PEOU) | 6 | 0.914 | 0.886 | 0.952 |
| Perceived Usefulness (PU) | 6 | 0.934 | 0.927 | 0.938 |
| Self-Efficacy (SE) | 10 | 0.937 | 0.944 | 0.917 |
| Intention to Use (IU) | 2 | 0.852 | 0.823 | 0.894 |

Construct validity examines whether the measurement scales represent and act like the attributes being measured. Principal component factor analysis with Varimax rotation was used to measure the construct validity. Table 3 shows that there are four factors. SE1 and SE2 together form a separate factor. These two items were subsequently removed from the analysis and we retained SE3 to SE10 as the SE factor. All the loadings were greater than 0.6, indicating strong loading on the factors determined through the Eigenvalue that is greater than 1.0 criterion. Total cumulative variances were 74.1% and 79.8% for males and females respectively.

Table 3. Factor loadings

| | Items | Male | Female |
|--------------------------------|--------------------------------|---------------|---------------|
| Factor 1 Perceived Usefulness | PU1 | 0.881 | 0.894 |
| | PU2 | 0.882 | 0.910 |
| | PU3 | 0.878 | 0.902 |
| | PU4 | 0.872 | 0.838 |
| | PU5 | 0.716 | 0.715 |
| | PU6 | 0.839 | 0.893 |
| Factor 2 Perceived Ease of Use | PEU1 | 0.687 | 0.686 |
| | PEU2 | 0.829 | 0.728 |
| | PEU3 | 0.733 | 0.804 |
| | PEU4 | 0.611 | 0.726 |
| | PEU5 | 0.737 | 0.821 |
| | PEU6 | 0.862 | 0.769 |
| Factor 3 Self-Efficacy - I | SE1 | 0.579 | 0.693 |
| | SE2 | 0.765 | 0.816 |
| Factor 4 Self-Efficacy - E | SE3 | 0.767 | 0.806 |
| | SE4 | 0.761 | 0.853 |
| | SE5 | 0.853 | 0.800 |
| | SE6 | 0.878 | 0.830 |
| | SE7 | 0.848 | 0.841 |
| | SE8 | 0.688 | 0.689 |
| | SE9 | 0.878 | 0.830 |
| | SE10 | 0.896 | 0.645 |
| | Total cumulative % of variance | 74.191 | 79.834 |

Discriminant validity is demonstrated if the measure is not a mirror of other variables. Table 4 shows the inter-correlation (off diagonal elements) between the constructs and the square roots of the average variance extracted (diagonal elements). The square roots of the average variance extracted were all greater than their shared variances. This confirms the discriminant validity of the constructs (Fornell & Larcker, 1981). Interestingly SE shows a negative but small and insignificant correlation between IU and PU for the male data. A negative relationship between SE and other TAM factors has previously been reported by Hasan (2006), Hasan and Ali (2006) and Chau

(2001). The correlation table begins to show the differences between male and female characteristics with respect to the relationships between TAM factors. These differences will be further confirmed later using the entire TAM model.

Table 4. Correlations between factors

| Variables | Mean | S.D. | IU | SE | PU | PEOU |
|------------------------------|----------------|----------------|---------------------|--------------------|------------------|------------------|
| Intention to Use (IU) | 4.06 (4.77) | 1.63 (1.36) | 0.922 (0.951) | | | |
| Self-Efficacy (SE) | 5.28 (5.09) | 1.24 (1.05) | -0.191 (.347*) | 0.843 (0.776) | | |
| Perceived Usefulness (PU) | 4.21 (4.79) | 1.36 (1.28) | 0.687** (.750**) | -0.146 (.290*) | 0.847 (0.861) | |
| Perceived Ease of Use (PEOU) | 5.26 (5.35) | 1.01 (1.11) | 0.224 (.590**) | .494** (.588**) | 0.21 (.599**) | 0.748 (0.757) |

Note. ** Pearson correlation is significant at the 0.01 level (1-tailed).

*Correlation is significant at the 0.05 level (1-tailed). Male (Female)

Diagonal elements are the average variance extracted.

Comparing Factor Scores

With reference to Table 5, the mean values on SMSEFF and SE were reported higher for males than females. This echoes previous research that males perceive themselves to have a higher skill base (Hartzel, 2003). Interestingly, females were reported to have higher mean values on OPAC experience, IU, PU and PEOU factors. However, only the factors PU and IU appear to differ significantly between males and females. It is possible that females see the system to be more work-related than males do (Ford & Miller, 1996).

Table 5. Individual scores and gender differences

| Items | Gender | N | SD | Mean | F | Sig. |
|------------------------------|--------|----|------|------|-------|---------------|
| OPAC [†] Experience | M | 54 | 0.82 | 2.33 | 0.821 | 0.367 |
| | F | 35 | 0.82 | 2.17 | | |
| SMS Efficiency | M | 54 | 1.43 | 5.80 | 2.080 | 0.153 |
| | F | 35 | 1.47 | 5.34 | | |
| Self-Efficacy | M | 55 | 1.24 | 5.28 | 0.525 | 0.471 |
| | F | 35 | 1.05 | 5.09 | | |
| Intention to Use | M | 55 | 1.63 | 4.06 | 4.573 | 0.035* |
| | F | 35 | 1.36 | 4.77 | | |
| Perceived Usefulness | M | 55 | 1.36 | 4.21 | 4.049 | 0.047* |
| | F | 35 | 1.28 | 4.79 | | |
| Perceived Ease of Use | M | 55 | 1.01 | 5.26 | 0.163 | 0.687 |
| | F | 35 | 1.11 | 5.35 | | |

Note. * P<0.05, † in reversed scale

Regression Model and Hypothesis Testing

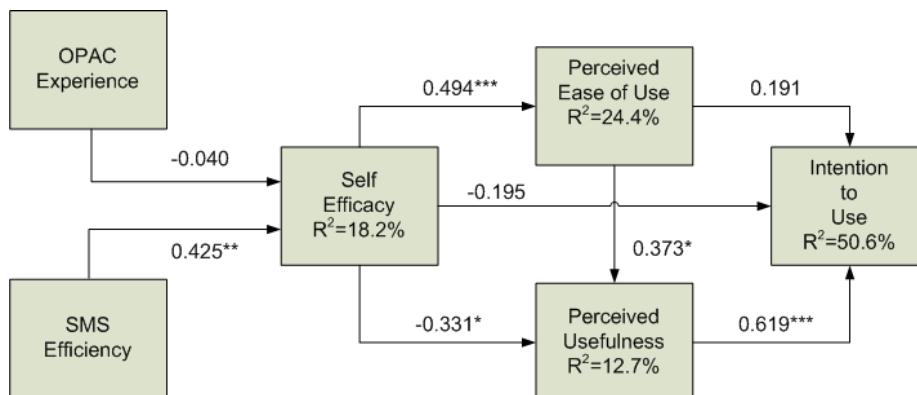
Regression analysis was performed for the male and female samples to generate the path coefficients. The individual path coefficients, significance, and variance explained are shown in Table 6, Figure 4, Table 7 and Figure 5 respectively. Subsequently, test procedures as recommended by Cohen and Cohen (1983) and Paternoster et al.

(1998) were used to measure any significant differences in the path coefficients between models. Table 8 shows that hypotheses H1, H2, H3, H4 and H5 were supported but not H6, H7 and H8.

Table 6. Male regression model

| Regression equation | R ² = | Beta | t-value | Significance |
|---------------------|------------------|----------|---------|--------------|
| SMSEFF+OPAC→SE | 0.182 | | | |
| SMSEFF | | 0.425** | 3.323 | 0.002 |
| OPAC [†] | | -0.040 | -0.316 | 0.753 |
| SE→PEOU | 0.244 | | | |
| SE | | 0.494*** | 4.137 | 0.000 |
| SE+PEOU→PU | 0.127 | | | |
| SE | | -0.331* | -2.218 | 0.031 |
| PEOU | | 0.373* | 2.501 | 0.016 |
| SE+PEOU+PU→IU | 0.506 | | | |
| SE | | -0.195 | -1.645 | 0.106 |
| PEOU | | 0.191 | 1.590 | 0.118 |
| PU | | 0.619*** | 5.876 | 0.000 |

Note. Standardized betas are reported, * P<0.05, ** P<0.01, *** P<0.001; † OPAC scale is reversed.



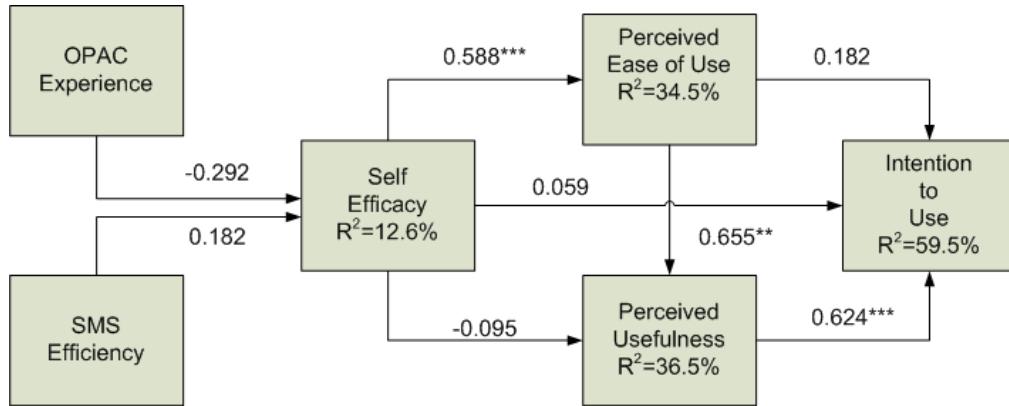
Note. Standardized betas are reported, * P<0.05, ** P<0.01, *** P<0.001

Figure 4. Male model

Table 7. Female regression model

| Regression equation | R ² = | Beta | t-value | Significance |
|---------------------|------------------|----------|---------|--------------|
| SMSEFF+OPAC→SE | 0.126 | | | |
| SMSEFF | | 0.182 | 1.099 | 0.280 |
| OPAC [†] | | -0.292 | -1.761 | 0.088 |
| SE→PEOU | 0.345 | | | |
| SE | | 0.588*** | 4.172 | 0.000 |
| SE+PEOU→PU | 0.365 | | | |
| SE | | -0.095 | -0.545 | 0.589 |
| PEOU | | 0.655** | 3.764 | 0.001 |
| SE+PEOU+PU→IU | 0.595 | | | |
| SE | | 0.059 | 0.418 | 0.679 |
| PEOU | | 0.182 | 1.071 | 0.293 |
| PU | | 0.624*** | 4.349 | 0.000 |

Note. Standardized betas are reported, ** P<0.01, *** P<0.001; † OPAC scale is reversed



Note. Standardized betas are reported, ** P<0.01, *** P<0.001

Figure 5. Female model

Table 8. Path Comparisons and Hypothesis Tests

| Relationship | B1 Male (n=55) | B2 Female(n=35) | Hypothesis | Difference between female and male | |
|--------------|-------------------|--------------------|-------------|--|----|
| SE-->IU | -0.255 | 0.077 | Female>male | †† | H1 |
| SE-->PU | -0.362 | -0.116 | Female>male | † | H2 |
| SE-->PEOU | 0.402 | 0.625 | Female>male | †† | H3 |
| SMSEFF-->SE | 0.369 | 0.129 | Male>female | †† | H4 |
| OPAC-->SE | -0.061 | -0.371 | Female>male | † | H5 |
| PEOU-->IU | 0.307 | 0.223 | Female>male | ns | H6 |
| PEOU-->PU | 0.502 | 0.755 | Female>male | ns | H7 |
| PU-->IU | 0.740 | 0.664 | Female>male | ns | H8 |

Note. Actual betas are reported † P<0.1, †† P<0.05

Discussion and Implications

From the results generated, the overall model has a better fit for females than males. The female model explains 59.5% of the variance whereas the male model explains 50.6% of the variance in IU. These numbers are comparable to previous SMS-TAM studies (Lee et al., 2007; Park & Chen, 2007; Turel et al., 2007; Pedersen, 2006; Mallat et al., 2006; Dennis & Venkatesh, 2003; Lai, 2004). Thus, the result supports the modified TAM model and explains the adoption of the proposed system satisfactorily. The test results reveal that all three antecedent constructs –SE, PU, and PEOU – have influence over IU. In both models, PU was shown to have the strongest direct impact on IU and there was no difference between females and males. H8 is thus not supported and this finding is different from that of Venkatesh and Morris (2000) but is similar to the findings of Nysveen et al. (2005) and Chiu et al. (2005). PEOU has insignificant direct impact on IU for both genders, thus H6 is not supported. This finding is again similar to that of Nysveen et al. (2005). H7 is also not supported as there is no difference between genders in the impact of PEOU on PU. This finding is similar to that of Venkatesh and Morris (2000). Possible explanations for the failure of these hypotheses are the similarity of the search system and familiarity with sending SMS which is already extremely popular among both genders. Also, treating gender as a psychological construct and not biological sex as suggested by Chui et al. (2005) in the current constructs and context may offer a plausible explanation.

However differences occur between males and females regarding the SE construct. It is observed that the factor has a negative and significant relationship with PU in the male model. This finding concurs with that of Chau (2001). One possible explanation for the negative relation is that the system is perceived to be very straightforward for experienced male SMS users and further external help as worded in the SE construct may result in users not appreciating the usefulness of the system. Chau (2001) suggested that users with high SE may see some limitations of a system in additional to its usefulness especially for non-sophisticated systems.

SE has no significant direct impact on IU in both models. The findings concur with those of Igbaria and Iivari (1995). However SE has a stronger positive impact on PU, PEOU and IU for females than males. Thus H1, H2 and

H3 are supported although H2 is only significant at the 0.1 level. The findings show that SE is perceived differently by males and females. The results show that the impact on SE through SMSEFF is significant for males but is not significant for females. It is likely that males and females perceived SE from different perspectives. Females may see the SMS based library search system as a new system that is less related to SMS usage. Therefore any attempts to improve SMSEFF as a proxy to improve SE for females may not be as effective in achieving the desired adoption objective. Other methods of enhancing females' SE, with the aim of demonstrating ease of use, would be beneficial.

The SE factor is shown to be more influenced by SMSEFF for males than females. Thus H4 is supported. On the other hand, OPAC experience has a higher impact on SE for females than males. Thus H5 is supported albeit at the 0.1 level. This shows that males tend to rely on the ability to "use" the technology to make adoption decisions and females tend to relate previous experience to current practices.

There are practical implications for educators and librarians who wish to promote the use of SMS based mobile catalog search to students. Based on our findings, it is recommended that they promote the usefulness of the system regardless of gender but put more emphasis on system usefulness for male users as they showed lower scores on PU. Using word-of-mouth strategy is recommended for male students to promote actual usage and usefulness as male students showed a significantly lower score on IU than female students. For female students, finding ways to increase their SE will be an effective solution to promote such a service. This could be achieved through special training on information literacy requirements combined with mobile search systems. Educators can design teaching and assignments suitable for mobile search scenarios to promote its usefulness. Based on the findings, increasing the level of experience with OPAC through additional training would be suitable for females as this will lead to stronger impact on SE and lead to higher impact on adoption. For male students, training should include the use of effective SMS texting skills such as predictive text and techniques to further increase SE and further emphasizing and demonstrating the ease of use of the system. System designers should take note that although the system is based on widely adopted SMS technology and assumed to be easy to use the adoption preference is still rather low especially with male students. There is a need to further identify the cause of low IU. Further analysis on the SE construct reveals that students preferred an online help and demonstration facility. Hence designers should design help facilities either in the form of a downloadable user manual or a key words dictionary and short demonstration video clips. Some form of personalization using a SMS template could also be implemented in this aspect.

Conclusion

As one of very few studies on gender differences in mobile search systems, this paper has provided insights into the different adoption patterns between genders based on a modified TAM which included constructs of PU, PEOU, IU, SE, SMSEFF and previous OPAC experience on a proposed SMS-based mobile library search system. The study confirms that gender differences still exist even for a technology that is already widely adopted among the younger generation. Therefore this research increases our current knowledge of gender differences in mobile search systems. The research helps practitioners and researchers better understand the different adoption characteristics between genders and thus how to train, educate and promote such a technology for supporting students' learning and improving information literacy skills. Continued research, development and evaluation is definitely required to provide further understanding about other potential factors that may have an impact on the acceptance of the proposed system and to provide useful guidelines for system designers and educators.

The current study has a number of limitations which could be addressed by further research. First, mobile search or mobile library search is a rather new and emerging technology and has yet to be widely considered among the usual information literacy skills. Therefore the findings here represent the younger population with limited mobile library search experience and should not be overly generalized. A similar study for seniors would be strongly recommended to develop different strategies (McCloskey, 2006). Second, the prototype is still evolving. It is anticipated that at different levels of complexity, the prototype system's user interfaces and features may have an influence regarding its usefulness. A well-developed design would be desirable in future studies and iterative evaluation may be necessary. Third, further studies on previous experience will be useful for implementing effective training. Besides OPAC experience there is a need to quantify some other possible measures of previous experience on other mobile search system experiences. Fourth, the sample size in this research is not large. This should be increased to a size that allows the use of the structure equation modeling technique to allow measurement error to be analyzed as an integral part of the model (Gefen et al., 2000). Fifth, the current study explores biological sex as a discriminating feature. It

will become important to identify individual psychological characteristics that fit into gender type rather than sex as a discriminating feature (Hupfer & Detlor, 2006). This could result in better understanding and effective strategies to support students' learning and academic success.

References

- Amin, H. (2007). Extending the technology acceptance model for SMS banking: Analyzing the gender gap among students. *International Journal of Business and Society*, 8(1), 15–27.
- Auckland Council (2010). *AirPAC – Your library in your hand*. Retrieved March 3, 2011 from <http://www.waitakere.govt.nz/CnlSer/libs/AirPac.asp>.
- Bandura, A. (1994). Self-efficacy. In Ramachaudran, V. S. (Ed.), *Encyclopedia of human behavior* (pp. 71-81). New York: Academic Press.
- Bradley, J., & Lee, C. C. (2007). ERP training and user satisfaction: A case study. *International Journal of Enterprise Information Systems*, 3(4), 33–50.
- Bundy, A. (2004). *Australian and New Zealand Information Literacy Framework: Principles, standards and practice* (2nd Ed.). Adelaide: ANZIIL.
- Candy, P. C., Crebert, G., & O'Leary, J. (1994). *Developing lifelong learners through undergraduate education*. Canberra: National Board of Employment, Education and Training, Australian Government Publishing Service.
- Chang, S. E. (2005). Computer anxiety and perception of task complexity in learning programming-related skills. *Computers in Human Behavior*, 21(5), 713–728.
- Chau, P. Y. K. (2001). Influence of computer attitude and self-efficacy on IT usage behaviour. *Journal of End User Computing*, 13(1), 26–33.
- Chiu, Y. B., Lin, C. P., & Tang, L. L. (2005). Gender differs: assessing a model of online purchase intentions in e-tail service. *International Journal of Service Industry Management*, 16(5), 416–435.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: development of a measure and initial test. *MIS Quarterly*, 19(2), 189–211.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Dennis, A. R., Venkatesh, V., & Ramesh, V. (2003). Adoption of collaboration technologies: Integrating technology acceptance and collaboration technology research. *Working Papers on Information Systems*, 3(8). Retrieved February 8, 2010, from <http://sproutsaisnet.org/3-8>.
- Ford, N., & Miller, D. (1996). Gender differences in internet perception and use. In Collier, M. & Arnold, K. (Eds.), *Electronic library and visual information research, Papers from the third ELVIRA conference*, (87–100), London: ASLIB.
- Fornell, C., & Larcker, D F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39–50.
- Fox, M. K. (2008). *PDAs, handhelds and mobile technologies in libraries – How the academic library is using handheld mobile technologies*. Retrieved February 4, 2009 from <http://web.simmons.edu/~fox/pda/>.
- Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the technology acceptance model, *MIS Quarterly*, 21(4), 389–400.
- Gefen, D., Straub, D., & Boudreau, M. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of AIS*, 4(7), 1–78.
- Giles, N., & Grey-Smith, S. (2005). TXTing librarians @ Curtin. *Paper presented at Information Online 2005*, 1–3 February, Sydney.
- Gorlenko, L., & Merrick, R. (2003). No wires attached: Usability challenges in the connected mobile world. *IBM Systems Journal*, 42(4), 639–651.

- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th Ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hartzel, K. (2003). How self-efficacy and gender issues affect software adoption and use. *Communications of the ACM*, 46(9), 167–171.
- Hasan, B. (2006). Delineating the effects of general and system-specific computer self-efficacy beliefs on IS acceptance. *Information and Management*, 43(5), 565–571.
- Hasan, B., & Ali, J. (2006). The impact of general and system-specific self-efficacy on computer training learning and reactions. *Academy of Information and Management Sciences Journal*, 9(1), 17–35.
- Herman, S. (2007). SMS reference: Keeping up with your clients. *The Electronic Library*, 25(4), 401–408.
- Hill, J. B., Hill, C. M., & Sherman, D. (2007). Text messaging in an academic library: Integrating SMS into digital reference. *The Reference Librarian*, 47(1), 17–29.
- Hupfer, M. E., & Detlor, B. (2006). Gender and web information seeking: a self-concept orientation model. *Journal of the American Society for Information Science and Technology*, 57(8), 1105–1115.
- Igbaria, M., & Iivari, J. (1995). The effects of self-efficacy on computer usage. *OMEGA International Journal of Management Science*, 23(6), 587–605.
- Jensen, B. (2010). SMS text reference comes of age: The My Info Quest collaborative. *The Reference Librarian*, 51 (4), 264–275.
- Kishore, R., Lee, J., & McLean, E. R. (2001). The role of personal innovativeness and self-efficacy in information technology acceptance: An extension of TAM with notions of risk. *Proceedings of the International Conference on Information Systems, ICIS 2001, December 16–19, New Orleans*. (pp. 469-474). Atlanta, GA: AIS
- Lai, T. L. (2004). Service quality and perceived value's impact on satisfaction, intention and usage of Short Message Service (SMS). *Information Systems Frontiers*, 6(4), 353–368.
- Lee, M. O. K., Cheung, C. D. M., & Chen, Z. (2007). Understanding user acceptance of multimedia messaging services: An empirical study. *Journal of the American Society for Information Science and Technology*, 58(13), 2066–2077.
- Lukasiewicz, A. (2007). Exploring the role of digital academic libraries: Changing student needs demand innovative service approach. *Library Review*, 56(9), 821–827.
- Macpherson, K. (2004). Undergraduate information literacy: A teaching framework. *Australian Academic & Research Libraries*, 35(3), 226–41.
- Madigan, E. M., Goodfellow, M., & Stone, J.A. (2007). Gender, perceptions, and reality: Technological literacy among first-year students. In *Proceedings of the 38th SIGCSE technical symposium on computer science education, March 7–11, Covington, KY*. (pp. 410-414). New York, NY: ACM.
- Mallat, N., Rossi, M., Tuunainen, V. K., & Öörmä, A. (2006). The impact of use situation and mobility on the acceptance of mobile ticketing services. In *Proceedings of the 39th Hawaii international conference on system sciences, Kauai, January 4–7, vol.2* (pp. 42b) Washington, DC: IEEE Computer Society.
- MessageNet (2008). *Monash University library management system with MessageNet Case study*. Retrieved November 21, 2010, from <http://www.messagenet.com.au/>.
- McCloskey, D. W. (2006). The importance of ease of use, usefulness, and trust to online consumers: An examination of the technology acceptance model with older consumers. *Journal of Organizational and End User Computing*, 18(3), 47–65.
- Mihalic, K., Reitberger, W., Obermair, C., & Tschelegi, M. (2005). Context in mobile HCI. In *Proceedings of the 7th international conference on human computer interaction with mobile devices & services, Salzburg, September 19–22, Mobile HCI '05, Vol. 111* (pp. 355-356). New York: ACM Press.
- Nysveen, H., Pedersen, P. E., & Thorbjørnsen, H. (2005). Explaining intention to use mobile chat services: moderating effects of gender, *Journal of Consumer Marketing*, 22(5), 247–256.
- Nokia (2010). *Nokia teams with UNESCO for m-learning in Africa*. Retrieved November 21, 2010, from <http://www.mobile-ent.biz/news/39228/Nokia-teams-with-UNESCO-for-m-learning-in-Africa>.
- Ong, C.-S., & Lai, J.-Y. (2006). Gender differences in perceptions and relationships among dominants of e-learning acceptance. *Computers in Human Behavior*, 22(5), 816–829.
- Park, Y., & Chen, J. V. (2007). Acceptance and adoption of the innovative use of smartphone. *Industrial Management & Data Systems*, 107(9), 1349–1365.

- Parker, A. (2007). SMS – Its use in the digital library. In Goh, D., Cao, T., Sølvberg, I., & Rasmussen, E. (Eds.) *Proceedings of the 10th international conference on Asian digital libraries: looking back 10 years and forging new frontiers (ICADL'07) Hanoi, Vietnam, December 10-13*, (pp. 387–390). Berlin: Springer.
- Pasanen, I., & Muhonen, A. (2002). Library in your pocket. *Paper presented at the International Association of Scientific and Technical University Libraries conference*, Retrieved February 4, 2009, from <http://www.iatul.org/conference/02/Muhonen.pdf>.
- Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. R. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36(4), 859–866.
- Pedersen, P. E. (2006). Instrumentality challenged: The adoption of a mobile parking service. In Ling, R. and Pedersen, P. E. (Eds.), *Mobile communications: Re-negotiation of the social sphere*, (pp. 373-388). London: Springer.
- Ramayah,T., & Aafaqi, B. (2004). Role of self-efficacy in e-library usage among students of a public university in Malaysia. *Malaysian Journal of Library & Information Science*, 9(1), 39-57.
- Roy, M., & Chi, M. T. H. (2003). Gender differences in patterns of searching the web. *Journal of Educational Computing Research*, 29, 335–348.
- Selwyn, N. (2007). Hi-tech=guy-tech? An exploration of undergraduate students' gendered perceptions of information and communication technologies. *Sex Roles*, 56, 525–36.
- Sohn, S., & Lee, D. (2007). Gender gap in the usage of mobile phone as digital multimedia device: The case of South Korea. *Paper presented at the annual meeting of the International Communication Association*, Retrieved February 4, 2009, from http://www.allacademic.com/meta/p171847_index.html.
- Turel, O., Serenko, A., & Bontis, N. (2007). User acceptance of wireless short messaging services: Deconstructing perceived value. *Information Management*, 44(1), 63–73.
- UN Foundation (2009). *Vital wave consulting. mHealth for development: The opportunity of mobile technology for healthcare in the developing world*. Washington, DC and Berkshire, UK: UN Foundation-Vodafone Foundation Partnership.
- Venkatesh, V., Speier, C., & Morris, M. G. (2002). User acceptance enablers in individual decision making about technology: Toward an integrated model. *Decision Sciences*, 33(2), 297–316.
- Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quarterly*, 24(1), 115–139.
- Waldman, M. (2003). Freshman's use of library electronic resources and self-efficacy, *Information Research*, 8(2), paper 150. Retrieved March 7, 2010, from <http://informationr.net/ir/8-2/paper150.htm>.
- Weiser, E. B. (2000). Gender differences in internet use patterns and internet application preferences: A two-sample comparison. *CyberPsychology and Behavior*, 3, 167–77.
- Wood, R. E., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology*, 56, 407–415.

Time-Quality Tradeoff of Waiting Strategies for Tutors to Retrieve Relevant Teaching Methods

Wen-Chung Shih¹, Shian-Shyong Tseng^{1*}, Che-Ching Yang² and Tyne Liang²

¹Department of Applied Informatics and Multimedia, Asia University, Taichung, 41354, Taiwan (R.O.C.) //

²Department of Computer Science, National Chiao Tung University, Hsinchu, 30010, Taiwan (R.O.C.) //

sstseng@asia.edu.tw

*Corresponding author

ABSTRACT

As more and more undergraduate students act as voluntary tutors to rural pupils after school, there is a growing need for a supporting environment to facilitate adaptive instruction. Among others, a teaching method retrieval system is intended to help tutors find relevant teaching methods for teaching a particular concept. However, teaching methods demanded by tutors might not be in well-organized forms or even unwritten. Thus, the technology of conventional information retrieval cannot be simply applied to retrieve teaching methods. Therefore, we propose to invite experts (teachers) to join the support environment and provide ad hoc teaching methods. Nevertheless, it is difficult for tutors to decide how long they will wait for replies from remote teachers, which is called a due time setting problem. Furthermore, a trade-off inherent in the problem is that an early reply does not necessarily mean a good one. Our idea is to propose an interactive approach to deal with the uncertainty during retrieving teaching methods. In this study, we define this problem and evaluate several waiting strategies. A prototype has been designed and implemented, and experiments have been conducted to evaluate the performance of the four waiting strategies. The results show that waiting strategies have a significant impact on retrieval performance.

Keywords

Teaching method retrieval, Due time setting, Volunteer tutor, Availability, Trustworthiness

Introduction

Encouraged by the concept of Service Learning (Bringle & Hatcher, 1996), more and more undergraduate students are willing to serve as voluntary tutors. The e-Tutors Project (<http://www.dsg.fju.edu.tw/>) supported by the Ministry of Education, Taiwan, provides remote tutoring service for pupils in rural areas. The preliminary results show that the learning performance of students has been improved. However, for this service to further assist rural students, the goal of adaptive learning has to be considered in a tutoring process. In fact, most of voluntary tutors are freshmen or sophomores from universities, who might not be professional educators. Hence, a tutor supporting mechanism is needed to facilitate adaptive tutoring. Among others, a teaching method retrieval system can help tutors to find effective teaching methods for pupils.

Traditional information retrieval focuses on static text resources. Instead, teaching method retrieval focuses on ad hoc teaching method discovery. That is to say, the teaching method to be found might have not been represented in a well-organized form, and might need further processing by experts. Also, the default users of the teaching method retrieval system are tutors, instead of tutees. Specifically, a tutee is assigned a Contact Tutor and then this Contact Tutor can get assistance by finding experts in the supporting environment, such as a social network. As shown in Figure 1, when a tutor cannot successfully make the pupil understand a concept, the tutor can ask for help from a social network consisting of supporting experts and teachers.

The due time setting problem is one of the important issues for teaching method retrieval in social networks. The primary reason is that a criterion is required for decision making when a submitted query is not responded. Particularly if the query is not replied to in the expected time, which can be estimated by log mining, the user will be confused. There are two main difficulties in the Due Time Setting problem. First, it involves not only technical issues, but also subjective human factors. For example, the system can calculate the estimated due time according to the information of availability, and provide it to the user as a suggestion. This is fine when the requested peer answers in time. However, how should the user act if the requested peer does not respond in time? When unexpected situations happen, one flexible solution is to resort to users' wills. Second, too little information about peers for decision making is available. A mechanism of peer information acquisition is needed for human decision makers.

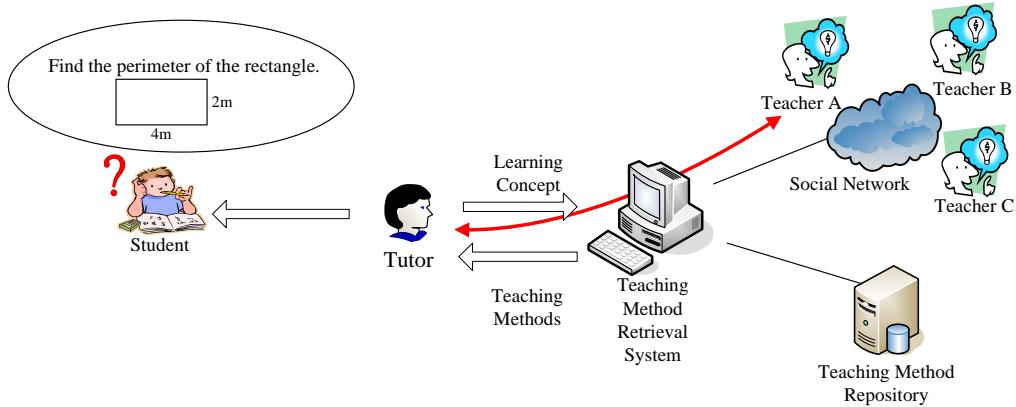


Figure 1. A tutor supporting environment

Our idea is to interactively set due time during the teaching method retrieval process. On one hand, tutors can interactively communicate with the experts to evaluate the time needed for getting a teaching method. On the other hand, tutors can estimate the time needed for getting a teaching method according to the reputation of experts. For example, some user might want to wait longer for a quality peer, even with low availability. We have compared several waiting strategies in this scenario. To realize this idea, a two-phase framework is proposed, which consists of a construction phase and a search phase. First, due time is set to a default value estimated by the system. When the query is due, the system asks the user whether to extend the due time or not. Other information is also provided for decision making.

Experiments are conducted to evaluate the performance of these waiting strategies. A tutor supporting environment is established by this prototype system, and 100 elementary-school teachers join this experiment to serve as supporting experts. The experiments investigate the importance of due time setting, the impact of supporting experts on performance and the waiting behavior of tutors. The results show that due time setting has significant impact on teaching method retrieval performance. Also, participants' opinions are reported.

The novelty of this paper lies in investigating the trade-off between the waiting time and quality of search results when retrieving teaching methods in social networks. The main idea is that users can interactively extend the waiting time to receive potentially desirable results. The main contribution of this work is the definition of the Due Time Setting problem for a tutor supporting environment, which addresses the trade-off between waiting time and performance of teaching method retrieval. The uniqueness of this problem lies in the integrated considerations of availability, trust and expert interests for teaching method retrieval. Furthermore, several due time setting strategies are proposed and the results of an evaluation are presented.

Purposes of the Study

The primary purpose of this study is to address the Due Time Setting problem which is one of the important issues for teaching method retrieval in social networks. Also, other purposes include designing and implementing a prototype for performance evaluation, and conducting experiments to evaluate the performance of the waiting strategies in the study.

- An empirical study was conducted to investigate the potential of this approach as a teaching methods retrieval tool. Three research questions were posed in this empirical study.
- i. What are the impacts of using different waiting strategies (Trustworthiness, Average Wait Time, Specialty and Static)?
 - ii. What are waiting behaviors of tutors?
 - iii. What are students' opinions on using this system?

Related Work

This work addresses the trade-off occurring in “indirect information retrieval”, which means retrieving ad hoc information (teaching methods) from experts, as shown in Figure 2.

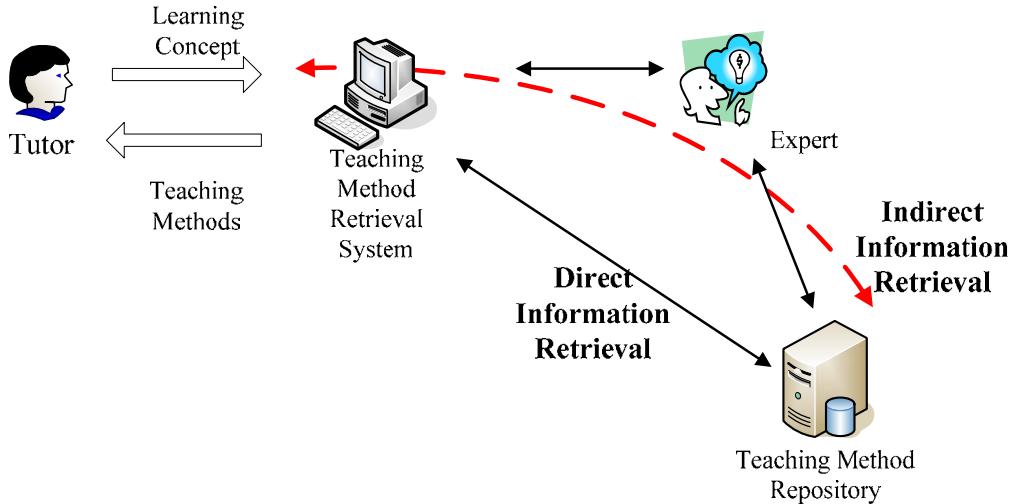


Figure 2. Indirect Information Retrieval

Information Retrieval

General information retrieval methods are mainly designed for web pages. Nevertheless, documents in specific domains may need tailor-made methods to improve their retrieval performance. For example, FAQ (Frequently Asked Questions) search (Kim et al., 2007; Kim & Seo, 2006; Wu et al., 2006) and patent retrieval (Fujita, 2007; Kang et al., 2007; Li & Shawe-Taylor, 2007) are widely investigated to find more efficient methods.

Search in peer-to-peer networks has been a flourishing research topic in recent years. A number of solutions to peer-to-peer search have been proposed (Nottelmann & Fischer, 2007; Parreira et al., 2007; Zhu et al., 2006). Zhu et al. (2006) indicated that one of the difficulties in peer-to-peer search is the lack of global statistical information, thus impeding the straightforward application of the well-known vector space model approach to peer-to-peer networks. Content search has been addressed in Shen et al. (2004) and Zeinalipour-Yazti et al. (2007). Nottelmann & Fischer (2007), Parreira et al. (2007) and Zhu et al. (2006) conducted information retrieval methodologies in peer-to-peer networks, while Yang & Chen (2008) studied social network.

Expert Finding

In recent years expert finding has been widely addressed in the field of information retrieval (Balog et al., 2009; Petkova & Croft, 2008; Zhu et al., 2009). A typical need for expert finding is as follows. An employee new to a company needs to know background knowledge on a project. An efficient and effective solution for the employee is to find an expert specializing in this project in the company. Existing models of expert finding assume that there is a large repository of documents. Then the task of expert finding is to identify experts specializing in a particular topic by analyzing the documents in the repository. Our work investigates finding experts from social networks, rather than from a large repository of documents. In the field of social networks, issues such as trustworthiness and availability are also important. The main reason for this is that: it is difficult to obtain useful expertise an expert who is not trustworthy or not available. Therefore, a feasible solution to retrieving teaching methods in a social network should consider these issues.

Wait Time Management

Issues related to waiting time setting have not been widely discussed in the literature. In Khan & Haque (2008), wait time management strategies are investigated in the context of P2P computing. In Stutzbach et al. (2006), the problem of sampling peer properties in peer-to-peer networks was addressed. They adopted an adaptive random walk approach to deal with departing peers. If a query times out, they try another peer from a stack. Tsui (2001) has reviewed techniques for peer-to-peer knowledge management. His e-learning applications have focused on the technologies of file sharing, distributed content networks and collaboration. To sum up, while related research tends to discover factors which affect the response time, we propose an interactive information retrieval environment where users can decide whether to wait for results, or not. With respect to conventional information retrieval, the waiting time has little impact on precision and recall of the retrieved results. However, in the proposed Tutor Supporting Environment it is possible for a tutor to wait longer for a teaching method, which otherwise could not be generated by the remote expert due to insufficient development time.

Problem Formulation

In this section, we formulate the Interactive Due Time Setting Problem for tutors to retrieve effective teaching methods in a social network. In a tutor supporting social network, a tutor issues a request for Teaching Methods (TM) about some Learning Concept. The problem is how the tutor can interactively set the due time to acquire an effective TM from a supporting teacher within a time constraint. The terms appearing in the problem formulation are defined as follows.

Notation and Assumption

This paper focuses on the teaching and learning “concepts” of a subject area. A “concept” is the basic unit for learning in a subject area. As shown in Figure 3, the concepts of the subject “Mathematics” are divided into four categories: Numbers, Geometry, Algebra and Probabilities. We assume that an ontology has been constructed to represent all concepts needed in this work. An ontology of learning concepts can be constructed by the method in Shih & Tseng (2009) and Shih et al. (2008a). Therefore, the task of a tutor is to make a pupil understand a “concept”.

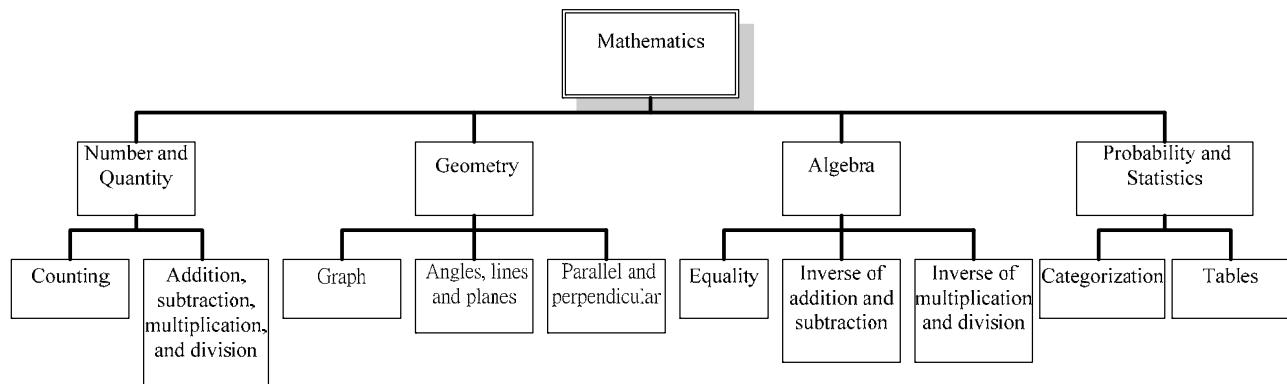


Figure 3. Ontology of Mathematics concepts

A Teaching Method (TM) is a sequence of operations, which can be followed by the tutor to help pupils understand the meaning of a concept. For example, a tutor finds that a pupil can not understand the concept of “the perimeter of a rectangle”. One Teaching Method offered by an experienced teacher might be “the demonstration method”, in which a building block is used by the tutor to indicate its perimeter. An Effective Teaching Method is a TM which can help the pupil to understand a concept.

In the interactive Tutor Support Environment, the feedback from remote experts plays an important role in due time setting. The due time estimated by the expert herself/himself is useful information for tutors’ reference. However, tutors need other objective information to evaluate the effectiveness of this estimate. Hence, the Expert Profile is

constructed and maintained to serve as the auxiliary information for decision making. In this paper, three of the main attributes of the supporting experts are characterized to analyze their impact on the performance of waiting strategies. In the following, the term *Transaction* is used to mean a communication session for a tutor to consult with an expert.

- This work addresses the trade-off between the waiting time and the quality (relevance) of a retrieved teaching method. However, there is another important attribute, trustworthiness, to be considered when retrieving teaching methods in social networks. For example, when a tutor asks an on-line expert to provide a teaching method, the tutor hopes that the expert can return a relevant teaching method as soon as possible. It is highly probable that the tutor's wish would come true if the expert is the tutor's "good friend" in the social network. That is to say, if the expert is a trustworthy friend of the tutor, the expert will make her/his best to provide a good teaching method as soon as possible. In this study trustworthiness (Trust) means the degree of friendliness, which is measured according to their historical transactions. When the expert promises to provide a teaching method and finally fails to reply, the Trust value is set to zero for this transaction. Otherwise, the Trust value of every transaction is set to one for this transaction. The Trust value of some expert is the average of Trust values in her/his historical transactions.

Policies for initializing the trustworthiness value of an expert depend on tutors. The optimistic policy can have 1 as the initial value, which means that the tutor assumes the expert a good friend. The pessimistic policy sets the initial value to 0, which means the tutor does not trust the expert. With a neutral policy, the initial value is 0.5. Updating of trust information occurs after each retrieval transaction. Therefore, the trustworthiness value of an expert will constantly evolve.

- Actual Waiting Time (*AWT*): The *AWT* value means the actual waiting time made by the expert in some transaction. When the expert promises to provide a *TM* and finally fails to reply, this transaction is not adopted for calculating the average *AWT* value. Therefore, the average *AWT* value is defined based on those transactions where the tutor provided a *TM* within the time limit.
- Specialty (*Specialty*): The *Specialty* value measures the proportion of providing relative teaching methods in each concept for some expert (supporting teacher). The relevance of a teaching method is judged by domain experts. The *Specialty* value in each concept for some expert is defined as follows.
-

$$\text{Specialty} = \frac{\# \text{relevant}}{\# \text{provided}} \quad (1)$$

Based on the three attributes, corresponding waiting strategies are designed and evaluated.

Problem Definition

In a common scenario of information retrieval, such as using a search engine, users input keywords as queries to search engines for retrieving relevant web pages. Search engines then evaluate the similarity between the keywords and web pages. Next, web pages relevant to the query are retrieved and ranked according to similarity computation. On the other hand, retrieving teaching methods in social networks is conducted in a peer-to-peer manner. For example, tutors identify experts in the social network, try to access the experts and then ask the on-line experts to provide their teaching methods.

In interactive tutor supporting environments, due time setting plays an important role in retrieving teaching methods. Therefore, this work is devoted to investigating the impact of waiting strategies on the performance of teaching method retrieval, which means that a tutor issues a request for Teaching Methods (*TM*) about some Concept in a tutor supporting environment. The problem is, given the profile of supporting experts and the admission of interaction with supporting experts, what is the impact of different waiting strategies on the performance for Teaching Method Retrieval. Especially, we focus on three waiting strategies according to the expert profile.

We use two criteria to evaluate the performance of waiting strategies. While Reduced Recall (RR) extends the classical Recall measure by considering the waiting time, Reduced Precision (RP) extends the classical Precision measure by considering the waiting time. The two criteria are defined as follows.

$$\text{Precision} = \frac{\# \text{relevant}}{\# \text{retrieved}} \quad (2)$$

$$Reduced_Precision = \frac{Precision}{(T + T_0)} \quad (3)$$

$$Recall = \frac{\#retrieved}{\#\text{responsive}} \quad (4)$$

$$Reduced_Recall = \frac{Recall}{(T + T_0)} \quad (5)$$

where

- $\#\text{relevant}$ is the number of relevant teaching methods;
- $\#\text{retrieved}$ is the number of retrieved teaching methods;
- $\#\text{responsive}$ is the number of experts who are willing to respond;
- T is the actual waiting time;
- T_0 is a constant representing the time taken to initialize the transaction. In this paper, the T_0 value is set 0 since we adopt relatively larger due time T .

Interactive Waiting Strategies

The flow of Interactive Due Time Setting is an iterative process, as shown in Figure 4. Each of the iterations is composed of the following steps:

- Set Initial Due Time. When context information is not available, the initial due time can only be roughly set. However, in the interactive tutor supporting environment, the feedback from the remote expert is a good value for the initial due time.
- Set Waiting Strategy. According to available context information, such as Trustworthiness, various waiting strategies can be adopted to set a good value for the due time.
- Calculate Due Time. This step calculates the due time according to the initial due time and the waiting strategy derived in the previous steps.
- Wait for Response. When the due time is determined, the tutor waits for replies from remote experts. If the tutor duly receives teaching methods, the process ends. Otherwise, the tutor can reset the due time by initiating a new iteration.

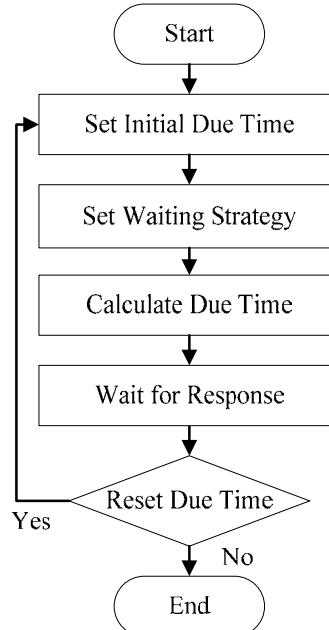


Figure 4. The flow of Interactive Due Time Setting

We have proposed a preliminary classification of plausible waiting strategies for the purpose of evaluation, as shown in Figure 5. These waiting strategies can be classified into three categories according to available context information: Trustworthiness, Actual Waiting Time and Specialty. In general, the setting of due time can be represented by the following formula:

$$T_{Due} = w(Trust, AWT, Specialty) \times T_{Initial} \quad (6)$$

where

- T_{Due} is the due time derived from the waiting strategy;
- $T_{Initial}$ is the initial value of the due time set by the tutor;
- w is the weighting function for adjusting the initial due time. This function has three parameters: *Trust*, *AWT* and *Specialty*. A waiting strategy can employ a combination of the three parameters to derive its due time.

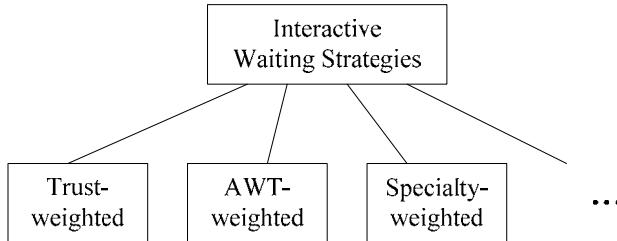


Figure 5. Classification of Interactive waiting strategies

The proposed waiting strategy of a tutor can be modeled by the weighting function in Formula (6) which is a combination of three parameters: *Trust*, *AWT* and *Specialty*. A tutor can work out her/his waiting strategy according to experts' trustworthiness, average response time and specialty. In this preliminary study we address three basic and independent strategies. The meanings of these strategies are explained as follows.

Trust-weighted Strategies

With *Trust*-weighted strategies the due time is determined mainly according to experts' trustworthiness. The more trustworthy the expert is, the longer waiting time will be derived. The physical meaning behind this strategy is to trade off additional time for possible quality content. Therefore, we expect to adjust the initial due time by the following formula:

$$T_{Due} = \alpha_{Trust} \times Trust \times T_{Initial} \quad (7)$$

where

- *Trust* is the historical trustworthiness value of the expert ($0 \leq Trust \leq 1$);
- α_{Trust} is a constant larger than one, which controls the range of the derived due time.

AWT-weighted Strategies

The historical behavior of the expert can also be useful information for adjusting the initial due time. If the *AWT* (Actual Waiting Time) value of the expert is larger than the waiting time claimed by the expert, we expect to increase the initial due time by the following formula:

$$T_{Due} = \alpha_{AWT} \times AWT \times T_{Initial} \quad (8)$$

where

- *AWT* is the historical value of Actual Waiting Time of the expert;
- α_{AWT} is a constant, which controls the range of the derived due time.

Specialty-weighted Strategies

With the goal of solving the question of pupils, the tutor expects to find a good teaching method, instead of a quick reply. Therefore, it is reasonable to extend the initial due time when the expert has a high *Specialty* value in the

learning concept. If the *Specialty* value of the expert is high ($0 \leq \text{Specialty} \leq 1$), the waiting time given by the expert would be extended to increase the possibility of obtaining a reply. Therefore, we adjust the initial due time by the following formula:

$$T_{\text{Due}} = \alpha_{\text{Specialty}} \times \text{Specialty} \times T_{\text{Initial}} \quad (9)$$

where

- *Specialty* is the historical Specialty value of the expert;
- $\alpha_{\text{Specialty}}$ is a constant, which controls the range of the derived due time.

Prototype: A Two-Phase Tutor Supporting System

To realize a Tutor Supporting Environment, a two-phase architecture is proposed, as shown in Figure 6. In the Search Phase, the tutor can retrieve teaching methods. The Interactive Teaching Method Retrieval Module enables the tutor to retrieve desirable teaching methods from the social network by submitting a query of some concept. In the Construction Phase, an Ontology is constructed and maintained to represent the concepts in the subject areas. In addition, an Expert Profile is constructed and maintained to collect and manage information about experts for decision making.

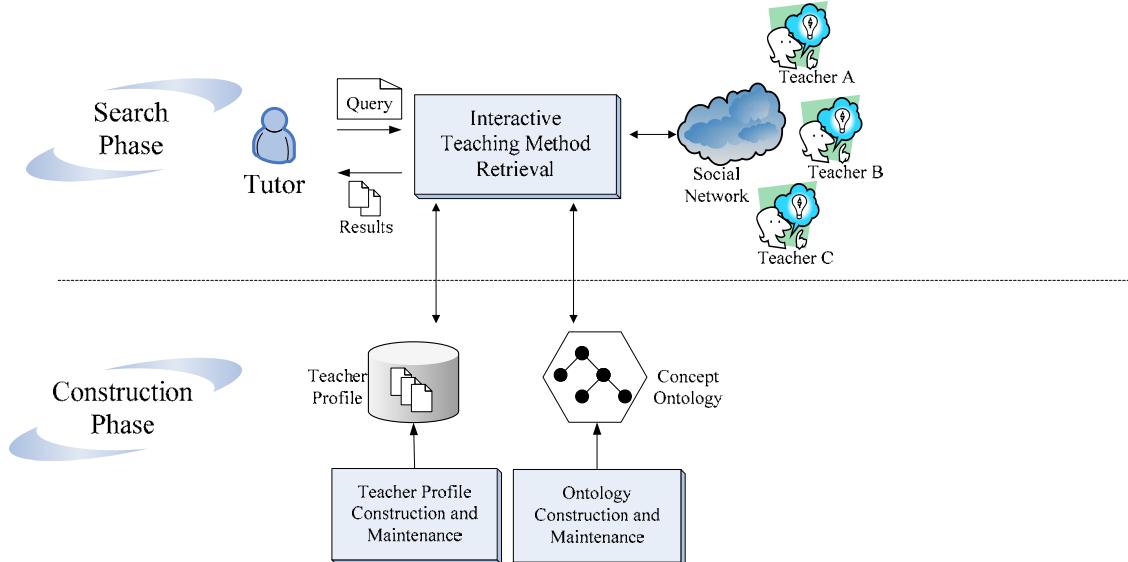


Figure 6. Overview of the interactive teaching method retrieval

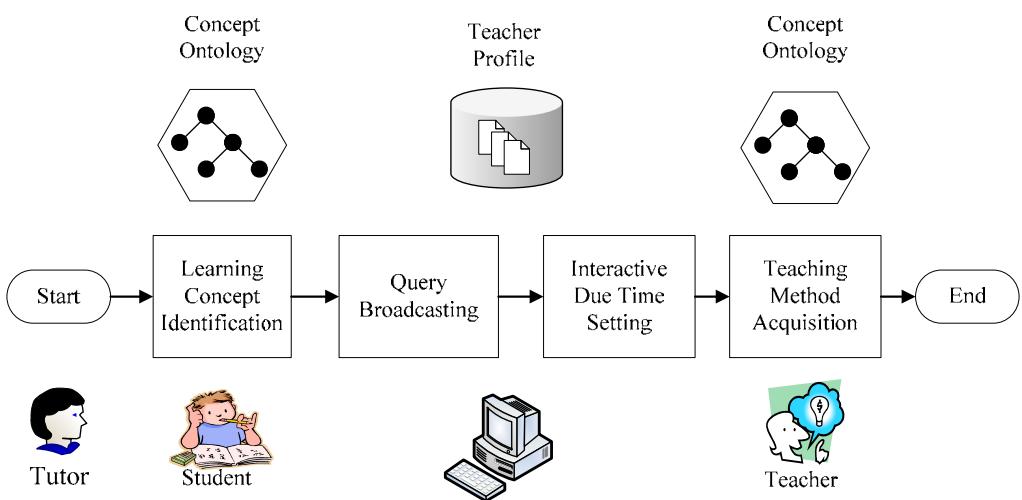


Figure 7. The flow of Search Phase

The flow of the aforementioned Interactive Teaching Method Retrieval Module is shown in Figure 7. The stage of Due Time Setting can be implemented with various waiting strategies. We have implemented the strategies for evaluation.

The initial ontology is constructed based on the Mathematics Ontology, as shown in Figure 3, by the 100 participating teachers. This ontology includes 100 concepts, which are equally divided into four categories: Numbers, Geometry, Algebra and Probabilities. With the built ontology, Expert Profile can be represented with respect to each of the concepts. In this implementation, the expert's profile includes Trust, AWT and Specialty. An example is shown in Figure 8. For the concept "C₂", this teacher has a high *Trust* value of 0.8. The average waiting time for "C₂" teaching method is 7 minutes. Also, the teacher has a middle degree of specialty, 0.5.



| Teacher | | C ₁ | C ₂ | C ₃ | ... | C _n |
|-----------|-----|----------------|----------------|----------------|-----|----------------|
| Trust | 0 | 0.8 | 0.7 | ... | 0.4 | |
| AWT | 20 | 7 | 16 | ... | 13 | |
| Specialty | 0.1 | 0.5 | 0.3 | ... | 0.6 | |

Figure 8. An example of Expert Profile of some teacher

Experimental Results

In this study, there are 100 pupils, 100 undergraduate tutors and 100 elementary-school teachers who volunteer to participate in this experiment and live in Taiwan. Each tutor is assigned a pupil. All teachers are involved in the three experiments. In the first experiment of Due time setting on static strategy, the participants are 4 pupils, 4 tutors and 100 teachers. In the experiment of waiting strategies evaluation, the participants are 10 pupils, 10 tutors and 100 teachers. In the last experiment of waiting behaviors of tutors, the participants are 100 pupils, 100 tutors and 100 teachers.

Experimental Setup

We have built a simulated Tutor Supporting Environment with the prototype system. A supporting community consisting of 100 elementary-school teachers was formed to help tutors via the Internet. To construct the profile of each supporting teacher, a training phase has been conducted for one month (from May 15 to June 15, 2009). On average, 25.5 supporting sessions have been collected for each teacher, and the distribution is shown in Figure 9.

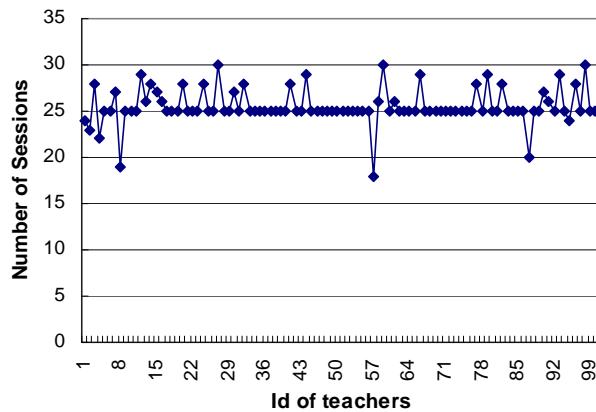


Figure 9. The distribution of supporting sessions

Two criteria have been adopted to evaluate the performance of waiting strategy. Reduced Precision (RP) extends the classical Precision measure by considering the waiting time. To calculate RP, it is necessary to determine the relevance of teaching methods to concepts. In this evaluation, the work of determining the relevance was executed by human experts. Particularly, five professors, experts in e-Learning, were invited to judge whether a teaching method was relevant to a Mathematical concept or not, the relevance being decided by vote.

Impact of Due Time Setting on Static Strategy

The purpose of this first experiment was to determine the importance of Due Time Setting for retrieving teaching methods in a tutor supporting environment. A Static waiting strategy was implemented as a baseline for evaluation. The Static strategy waits for a fixed amount of time. The participants are 4 pupils, 4 tutors and 100 supporting teachers. Each tutor is assigned a pupil. Without loss of generality, we let the 4 pupils study the four subjects: Numbers, Geometry, Algebra and Probabilities. Tutors use the tutor supporting environment to retrieve teaching methods. Each pupil asks three questions about learning concepts, and the average Precision and Recall values are derived. Figure 10 shows the results with respect to classical recall and precision. First, we observe that the Recall value increases as Due Time increases for the four subjects. This means that an adaptive waiting strategy is demanded and a naïve static waiting strategy will not be suitable for this scenario. Similarly, the Precision value increases as Due Time increases for the four subjects, though not so significantly. Furthermore, we find that Due Time Setting is suitable for the scenario of Recall-oriented teaching method retrieval, which tries to find as many results as possible. For Precision-oriented teaching method retrieval, longer waiting might not significantly increase the performance.

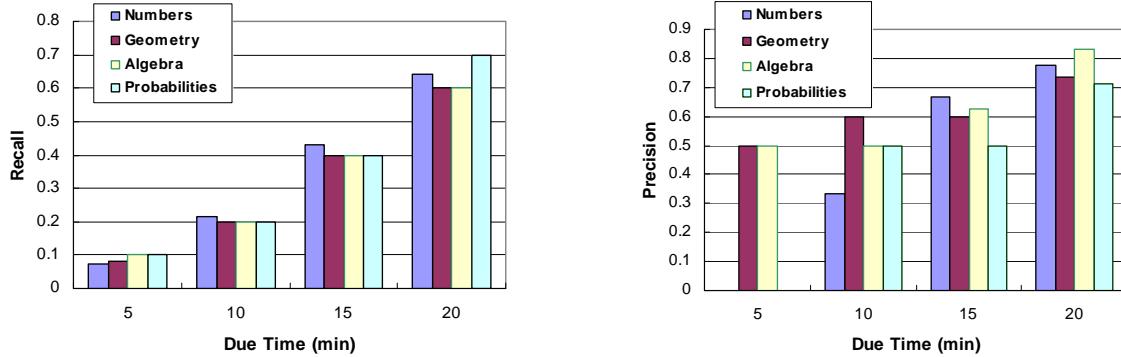


Figure 10. (a) The relation of Recall and due time; (b) the relation of Precision and due time

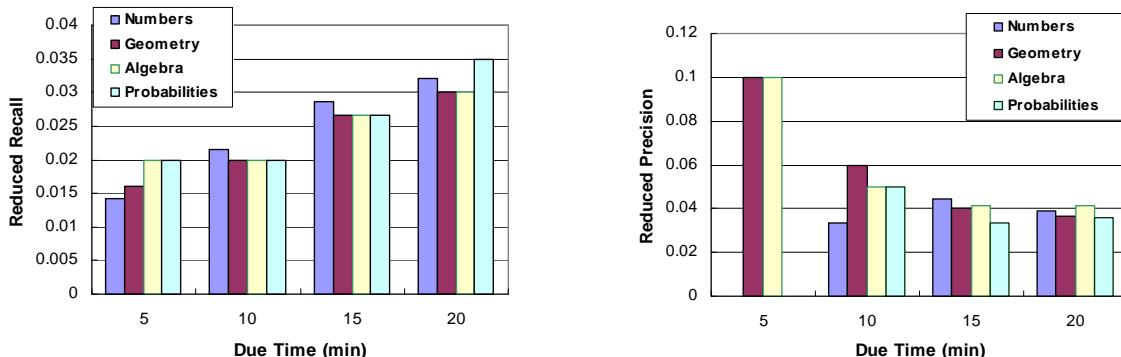


Figure 11. (a) The relation of Reduced Recall and due time; (b) the relation of Reduced Precision and due time

Next, since the classic Recall and Precision measures are related to the Due Time, a new metric is needed to evaluate the performance of various waiting strategies. We use the Reduced Recall and the Reduced Precision measures to present the results of Figure 10, as shown in Figure 11. We found that the two reduced measures take Due Time into

consideration for the performance of teaching method retrieval. Consequently, the performance does not depend on Due Time. The range of Reduced Recall falls between 0.01 and 0.035; the range of Reduced Precision falls between 0.03 and 0.06 (excluding the values for “Due Time = 5 minutes”). Note that, although the variance looks obvious in this figure, the range of variance is really small and can nearly be neglected.

Evaluation of Waiting Strategies

The next experiment was to understand the impact of the expert profile on performance, with respect to Trustworthiness, Average Wait Time and Specialty. A Static waiting strategy was implemented as a baseline for evaluation. The Static strategy waits for a fixed amount of time, which is 10 minutes in this experiment. We chose the value of 10 minutes because this is the common waiting time for the participating tutors, as presented in the next section. The participants were 10 pupils, 10 tutors and 100 supporting teachers. Each tutor was assigned a pupil. The profiles of the 100 teachers are shown in Tables 2-4. Table 2 shows the teachers divided into three groups according to their *Trust* values. For example, the teachers of “High Trust” group are highly trustworthy, which means they used to provide useful support according historical transaction records. Tutors can trust the experts and wait for their support. However, tutors should not count on the “Low Trust” experts even though the experts have high *Specialty* values or short waiting time.

Table 2. Groups of teachers for Trustworthiness

| Group Name | Trustworthiness Range | Number of Teachers |
|------------|--|--------------------|
| High Trust | $0.75 < \text{Trustworthiness} \leq 1$ | 45 |
| Mid. Trust | $0.5 < \text{Trustworthiness} \leq 0.75$ | 38 |
| Low Trust | $0 \leq \text{Trustworthiness} \leq 0.5$ | 27 |

Table 3 shows the teachers divided into three groups according to their *Average Wait Time* values. For example, the teachers of “Long Wait” group let tutors wait more than 15 minutes on average. If the waiting time estimated by the teacher is far from the *Average Wait Time*, the Due Time might need to be adjusted to avoid unnecessary waiting.

Table 3. Groups of teachers for Average Wait Time

| Group Name | Range of Average Wait Time | Number of Teachers |
|------------|--|--------------------|
| Long Wait | $15 \text{ min} < \text{Average Wait Time}$ | 23 |
| | $5 < \text{Average Wait Time} \leq 15 \text{ min}$ | |
| Mid. Wait | | 46 |
| Short Wait | $0 \leq \text{Average Wait Time} \leq 5 \text{ min}$ | 31 |

Similarly, Table 4 shows the teachers divided into three groups according to their *Specialty* values. For example, the teachers of “High Specialty” group provided relevant teaching methods, implying they are “experts” in those concepts. Therefore, we might be willing to wait longer for the “High Specialty” group to get useful teaching methods.

Table 4. Groups of teachers for Specialty

| Group Name | Specialty Range | Number of Teachers |
|----------------|------------------------------------|--------------------|
| High Specialty | $0.75 < \text{Specialty} \leq 1$ | 26 |
| Mid. Specialty | $0.5 < \text{Specialty} \leq 0.75$ | 53 |
| Low Specialty | $0 \leq \text{Specialty} \leq 0.5$ | 21 |

Figure 12 shows the results with respect to Reduced Recall and Reduced Precision, which are the average values of the 10 tutoring results. In general, we observe that the three strategies had a better performance than the Static strategy. This is reasonable because a fixed due time is not suitable for the Tutor Supporting Environment, where longer waiting time will result in better precision and recall.

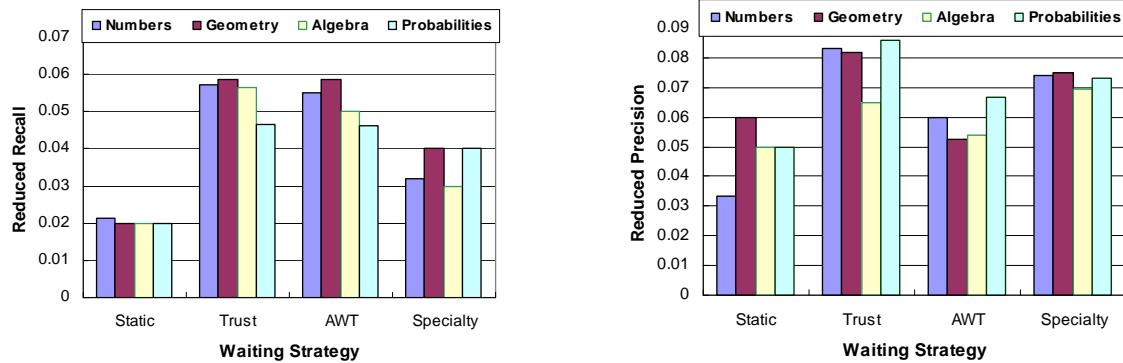


Figure 12. (a) Reduced Recall of waiting strategies; (b) Reduced Precision of waiting strategies

For Reduced Recall, as shown in Figure 12(a), the *Trust* strategy and the *AWT* strategy perform better than the *Specialty* strategy. The main reason might be that the tutors using the *Specialty* strategy tend to set longer due time than needed, which results in performance degradation. For Reduced Precision, as shown in Figure 12(b), the *Trust* strategy and the *Specialty* strategy had better performance. Based on this result, we find that the two strategies are helpful for finding relevant teaching methods.

Waiting Behavior of Tutors: Trade-off

The next experiment addressed the waiting behavior of the tutors. The participants were 100 pupils, 100 tutors and 100 supporting teachers. Each tutor was assigned a pupil. After usage training for one month, this prototype was used to form a tutor supporting environment and retrieve teaching methods. The waiting time of the 100 tutors are illustrated in Figure 13. We observed that most of the tutors set their due time to 10 minutes though larger due time results in better performance. Also, Figure 13 shows the results with respect to classic Precision and Recall, which is the average value of the 100 tutoring results.

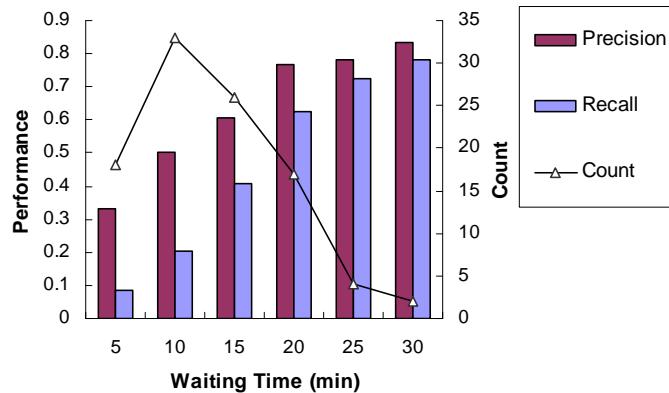


Figure 13. Waiting behavior of the tutors

Opinions of Students

To collect participants' opinions on this approach, we have conducted an open-ended-question interview with ten tutor participants and ten student participants. The open-ended questions are listed as follows.

- What is your main problem while teaching/learning mathematics?
- What do you think of using this system as an auxiliary teaching mechanism?

Tutors' opinions are summarized as follows.

- Certain mathematical concepts are difficult for students to understand. When students cannot fully understand after the instructions, one of the tutors said to students: "Practice more exercises and you will understand it someday." The other tutors said, "Just memorize the formulas." These tutors wanted to help students understand these concepts but they did not know the appropriate teaching methods. Oppositely, other tutors will look for useful teaching materials such as web searching or peer assistance.
- The tutors stated that the proposed approach can help them retrieve useful and valuable teaching methods. One of the tutors said, "Sometimes I do not know how to teach, this system can help me call for assistance." Another tutor said, "Using this system can retrieve different teaching methods toward this same topic. By using those teaching methods, my student can actually understand mathematical concepts." One tutor said, "The Trust strategy is more useful than the static strategy. With the aid of the Trust strategy, I can retrieve useful teaching materials within a time constraint." In addition, the suggestions of tutors are that the speed of communication mechanism need to be improved especially the network bandwidth and computer performance.

Students' opinions are summarized below, respectively.

- Some of the students stated mathematics is abstract and difficult to learn. When they could not understand mathematical concepts after they asked for help, one of the students said, "I will learn it by rote." The other student said, "I will give up." Consequently, it is difficult for them to finish their homework themselves.
- With the proposed approach tutors try alternative teaching methods to make them understand the meaning of mathematical concepts. Students will not hear such answers as "you will understand it someday" or "just memorize the formulas."

Discussion

As a consequence of these experiments we have several findings in response to the research questions. First, waiting strategies make a significant impact on the performance of teaching method retrieval in social networks. It is found in the first two experiments that the static waiting strategy results in poor performance. When the due time is set according to experts' attributes, such as trustworthiness, the performance of teaching method retrieval can be improved obviously. Among these strategies, the trust-based strategy had the best performance in the interactive tutor supporting environment, especially for Recall-oriented scenarios. Second, we found that more than half of the tutors wait for ten to fifteen minutes. Their behavior does not result in optimal performance. In fact, setting suitable due time will improve the retrieval performance. Finally, we found that the opinions of tutors and students are in favor of using the system.

Shih et al. (2008b) pointed out the benefits of teaching material sharing, which can effectively reduce the teaching loads of teachers. However, with the rapid growth of information, it is difficult to find the target information on the Internet (Saito & Miwa, 2007). Furthermore, tutors may not be familiar with the skills of information seeking. To assist tutors in retrieving target information, time-quality tradeoff of waiting strategies is proposed in this study. The experimental results show that the proposed three waiting strategies got better performance than static strategy. Tutors not only retrieve relevant teaching methods easily but also shorten his/her waiting time. Besides, tutors can get runtime assistance to retrieve relevant teaching methods and save time for assisting weak pupils when coaching them. As for the students, different thinking styles affect learning preferences (Sternberg, 1997). Furthermore, research has shown that the sooner students receive feedback the more effective it is for their learning (Irons, 2008). Students can benefit from feedback if they receive it before they move onto their next assignment. Based on the students' interview results, they have a positive attitude toward the proposed approach and their learning interests are aroused by using their preferred learning approach.

Trust computation mechanisms are important to P2P networks and social networks, which are used to update the degree to which a peer trusts other peers. The trust computation is similar to the weight learning appearing in many applications. For example, self-organizing maps of neural networks use similar, however more complex, methods to update neurons' weighting vectors. Trust computation in P2P networks has received great attention in recent years (Li & Ling, 2004; Ming et al., 2005; Zhou & Hwang, 2007), and many existing methods can be used to support the proposed due time setting method, with minor modifications. To highlight the proposed idea, we have defined the Trustworthiness of experts and calculated the Trust value by straightforward statistical methods.

In this work, both recall-oriented and precision-oriented measures are used. On one hand, the teaching tasks of the participants can be recall-oriented, such as educational content review. On the other hand, precision-oriented measures are suitable for question answering, where one correct document is enough. Furthermore, the calculation of precision involves the decision of relevance. In this work, relevance is decided by vote among five experts. However, the concept of Web 2.0 can be applied to the determination of relevance. In other words, relevance can be decided by all members in the Tutor Supporting Environment, including tutors, teachers and pupils.

Conclusions and Future Work

In this study, a tutor supporting environment is proposed and shown as a promising platform for remote tutoring services for pupils in rural areas. To reduce response time without sacrificing the quality of retrieved teaching methods, we formulate the due time setting problem in tutor supporting scenarios, and evaluate the performance of different waiting strategies.

Experimental results show that the dynamic nature of a social network will degrade the performance of retrieval strategies if they do not take the expert profile into account. Advantages of the proposed approach can be summarized as follows. First, the due time setting mechanism improves response time without sacrificing much recall. Based on the effective management of trust, average wait time and specialty information, the dynamic status of experts can be estimated. Second, the due time setting in an interactive manner redeems the possible failure in statistical estimation, and provides a flexible way for human users to make decisions of due time extension.

From the discussion above we see some educational implications of using waiting strategies in social network-based teaching method retrieval. As indicated by Yang et al. (2007), one of the important elements to achieve knowledge sharing in a virtual learning community is mutual trust. It is possible that an expert refuses to offer a teaching method to a tutor. Our study complements Yang et al.'s work by addressing the trade-off of waiting time and content quality and indicating the importance of trustworthiness for retrieving teaching methods. According to our findings, tutors should consider suitable waiting strategies in order to enhance the quality of retrieved teaching methods. In addition, it is desirable for tutors to identify trustworthy experts through historical transactions and then to enhance their mutual trust constantly. Our findings suggest a two-stage retrieval mechanism: maintain your own community of experts for consultation, and then set suitable waiting time to retrieve quality teaching methods.

In the near future, we plan to extend this work to investigate interaction between humans and systems in the tutor supporting environment. Furthermore, we will try to integrate the tutoring platform with a Wiki-based teaching material design.

Acknowledgements

This research was partially supported by National Science Council of Republic of China under the numbers: NSC97-2511-S-468-004-MY3, NSC98-2511-S-468-004-MY3, NSC98-2511-S-468-002 and NSC99-2511-S-468-003.

References

- Balog, K., Azzopardi, L., & de Rijke, M. (2009). A language modeling framework for expert finding. *Information Processing & Management*, 45(1), 1-19.
- Bringle, R. G., & Hatcher, J. A. (1996). Implementing Service Learning in Higher Education. *Journal of Higher Education*, 67 (2), retrieved May 1, 2011 from <http://www.compact.org/advancedtoolkit/pdf/bringle-all.pdf>.
- Fujita, S. (2007). Technology survey and invalidity search: A comparative study of different tasks for Japanese patent document retrieval. *Information Processing and Management*, 43(5), 1154-1172.
- Irons, A. (2008). *Enhancing learning through formative assessment and feedback*. Routledge, Abingdon, UK.
- Kang, I.-S., Na, S.-H., Kim, J., & Lee, J.-H. (2007). Cluster-based patent retrieval. *Information Processing and Management*, 43(5), 1173-1182.

- Khan, J. I., & Haque, A. U. (2008). Computing with data non-determinism: Wait time management for peer-to-peer systems. *Computer Communications*, 31(3), 629-642.
- Kim, H., Lee, H., & Seo, J. (2007). A reliable FAQ retrieval system using a query log classification technique based on latent semantic analysis. *Information Processing and Management*, 43(2), 420-430.
- Kim, H., & Seo, J. (2006). High-performance FAQ retrieval using an automatic clustering method of query logs. *Information Processing and Management*, 42(3), 650-661.
- Li, X., & Ling, L. (2004). PeerTrust: supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Transactions on Knowledge and Data Engineering*, 16(7), 843-857.
- Li, Y., & Shawe-Taylor, J. (2007). Advanced learning algorithms for cross-language patent retrieval and classification. *Information Processing and Management*, 43(5), 1183-1199.
- Ming, F., Yong, T., & Whinston, A. B. (2005). Evaluation and design of online cooperative feedback mechanisms for reputation management. *IEEE Transactions on Knowledge and Data Engineering*, 17(2), 244-254.
- Nottelmann, H., & Fischer, G. (2007). Search and browse services for heterogeneous collections with the peer-to-peer network Pepper. *Information Processing and Management*, 43(3), 624-642.
- Parreira, J. X., Michel, S., & Weikum, G. (2007). p2pDating: Real life inspired semantic overlay networks for Web search. *Information Processing and Management*, 43(3), 643-664.
- Petkova, D., & Croft, W. B. (2008). Hierarchical language models for expert finding in enterprise corpora. *International Journal on Artificial Intelligence Tools*, 17(1), 5-18.
- Saito, H., & Miwa, K. (2007). Construction of a learning environment supporting learners' reflection: A case of information seeking on the Web. *Computers & Education*, 49(2), 214-229.
- Shen, H. T., Shu, Y., & Yu, B. (2004). Efficient semantic-based content search in P2P network. *Transactions on Knowledge and Data Engineering*, 16(7), 813-826.
- Shih, W.-C., & Tseng, S.-S. (2009). A Knowledge-based Approach to Retrieving Teaching Materials for Context-aware Learning. *Educational Technology & Society*, 12(1), 82-106.
- Shih, W.-C., Tseng, S.-S., & Yang, C.-T. (2008a). Using Taxonomic Indexing Trees to Efficiently Retrieve SCORM-compliant Documents in e-Learning Grids. *Educational Technology & Society*, 11(2), 206-226.
- Shih, W.-C., Tseng, S.-S., & Yang, C.-T. (2008b). Wiki-based Rapid Prototyping for Teaching-Material Design in e-Learning Grids. *Computers & Education*, 51(3), 1037-1057.
- Sternberg, R. J. (1997). *Thinking Style*: New York: Cambridge University press.
- Stutzbach, D., Rejaie, R., Duffield, N., Sen, S., & Willinger, W. (2006). On unbiased sampling for unstructured peer-to-peer networks. *Paper presented at the Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*, Rio de Janeiro, Brazil.
- Tsui, E. (2001). Technologies for Personal and Peer-to-Peer (P2P) Knowledge Management. *Computer Sciences Corporation Leading Edge Forum*, 1-53.
- Wu, C.-H., Yeh, J.-F., & Lai, Y.-S. (2006). Semantic Segment Extraction and Matching for Internet FAQ Retrieval. *Transactions on Knowledge and Data Engineering*, 18(7), 930-940.
- Yang, S. J. H., & Chen, I. Y. L. (2008). A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network. *International Journal of Human - Computer Studies*, 66(1), 36-50.
- Yang, S. J. H., Chen, I. Y. L., Kinshuk, & Chen, N.-S. (2007). Enhancing the Quality of e-Learning in Virtual Learning Communities by Finding Quality Learning Content and Trustworthy Collaborators. *Educational Technology & Society*, 10(2), 84-95.
- Zeinalipour-Yazti, D., Kalogeraki, V., & Gunopoulos, D. (2007). pFusion: A P2P Architecture for Internet-Scale Content-Based Search and Retrieval. *Transactions on Parallel and Distributed Systems*, 18(6), 804-817.
- Zhou, R., & Hwang, K. (2007). Gossip-based Reputation Aggregation for Unstructured Peer-to-Peer Networks. *Paper presented at the IEEE International Parallel and Distributed Processing Symposium*, March 26-30, Long Beach, CA, USA.
- Zhu, J. H., Song, D. W., & Ruger, S. (2009). Integrating Multiple Windows and Document Features for Expert Finding. *Journal of the American Society for Information Science and Technology*, 60(4), 694-715.
- Zhu, X., Cao, H., & Yu, Y. (2006). SDQE: towards automatic semantic query optimization in P2P systems. *Information Processing and Management*, 42(1), 222-236.

Self-efficacy in Internet-based Learning Environments: A Literature Review

Chin-Chung Tsai¹, Shih-Chyueh Chuang², Jyh-Chong Liang² and Meng-Jung Tsai¹

¹Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan // ²Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan // cctsai@mail.ntust.edu.tw // fscchuang@tp.edu.tw // aljc@mail.ntust.edu.tw // mjtsai99@mail.ntust.edu.tw

ABSTRACT

This paper reviews 46 papers from 1999 to 2009 regarding self-efficacy in Internet-based learning environments, and discusses three major categories of research: (1) learners' Internet self-efficacy, assessing learners' confidence in their skills or knowledge of operating general Internet functions or applications in Internet-based learning; (2) the interplay between learners' general academic self-efficacy and their Internet-based learning, and (3) learners' self-efficacy, particularly in terms of Internet-based learning. In general, students' self-efficacy plays a positive role in their attitude towards and their processes and outcomes derived from Internet-based learning. It was found that for the reviewed studies, a significant amount of research has adopted search tasks to predict students' learning outcomes in Internet-based settings, implying that search tasks may still be considered as the most commonly implemented Internet-based learning activities. All of the studies utilized questionnaires or surveys for assessing students' self-efficacy, and mostly selected students in higher education institutes as their samples. It was also found that relatively few empirical studies were conducted from the theoretical perspectives of the initially proposed concept of self-efficacy.

Keywords

Internet, Self-efficacy, Internet-based learning, Internet self-efficacy

Introduction

In recent years, the Internet-based environment has experienced prolific development. With multiple forms of representation, some research has indicated that content delivered in blended or Internet-based conditions may be more efficient than that is provided in traditional classrooms (Abdous & Yoshimura, 2010). However, when comparing the differences between distance education and classroom-based instruction, some studies have found no significant difference in effectiveness between distance education within blended or Internet-based learning settings and face-to-face (F2F) education (Bernard *et al.*, 2004; Delialioglu & Yildirim, 2008). Indeed, a variety of results have been derived from the studies relating to Internet-based learning (IBL). For instance, one of the features of IBL is that it is supposed to foster learners' active participation in the construction of knowledge (White & Frederiksen, 2005). Some studies have indicated positive effects on students' knowledge construction in IBL processes (Pena-Shaff & Nicholls, 2004), whereas some have reported that the discussions among learners are generally at low levels of knowledge construction in IBL systems (Guan, Tsai & Hwang, 2006). Hence, it may be suggested that due to the divergent components of the IBL context created by a variety of researchers, different traits may be produced.

Undoubtedly, a large amount of work has been devoted to constructing a preferable Internet-based setting to date. When compared with the traditional classroom setting, IBL provides many appealing attributes, which may consist of increasing the availability of learning experiences for learners who cannot or choose not to attend F2F offerings, assembling and disseminating instructional content more cost-efficiently, or enabling instructors to handle more students while maintaining learning outcome quality that is similar to that of comparable F2F instruction (Abdous & Yoshimura, 2010). Furthermore, learners can have access to the information without time limits or location constraints. That is, Internet-based settings may help learning be unrestricted to any specific moment or to any particular classroom. Thus, it is generally believed that IBL is likely to provide potential applications for students' learning activities. However, some research indicates that users' intention to continue in IBL may be low (Lee, 2010). In other words, occurrences of participants dropping out of IBL are not uncommon (Roca, Chiu, & Martinez, 2006). A number of possible explanations may account for learners' discontinuing IBL. For instance, more than a decade ago, Katz and Aspeden (1996) stated that uncertainty about how to get started and the perception that computers were too complicated were possible barriers. Indeed, research evidence has indicated that learners are unwilling to take part in IBL probably because they lack confidence in operating the system (Eastin & LaRose, 2000).

Based on a similar thought, some researchers have suggested that students' involvement in IBL may be associated with the perception of their own capabilities relating to specific skills and knowledge. They assert that such a concept, often referred to as *self-efficacy*, may play an important role in students' learning processes and learning outcomes in Internet-based classroom settings (Shakpa & Ferrari, 2003). In fact, recent empirical studies (e.g. Hoffman & Spathariu, 2008) have demonstrated that students with higher self-efficacy gain better performance in contrast to those with lower self-efficacy in Internet-based settings. Thus, with the significant importance of self-efficacy in IBL, the aim of this paper is to conduct a literature review examining the theory, evidence, and application of the relationship between self-efficacy and IBL. On the basis of such a concept, the review will firstly define a theoretical framework for exploring self-efficacy in the Internet-based context. Then, the evidence of self-efficacy in IBL and how it is connected to the original concept of self-efficacy will be explained. Finally, a collection of related empirical studies regarding self-efficacy in the IBL condition will be reviewed. In this review, three categories regarding the relations between self-efficacy and IBL are classified:

- 1) the *Internet Self-Efficacy* (ISE), which examines learners' confidence in their general skills or knowledge of operating Internet functions or applications in the Internet-based learning condition;
- 2) the interplay between *Academic Self-Efficacy and Internet-Based Learning* (ASE&IBL), which investigates the role of learners' general academic confidence played in the Internet-based learning condition;
- 3) the *Internet-Based Learning Self-Efficacy* (IBLSE), which explores learners' confidence in their participation and their expected performance, particularly in terms of the Internet-based learning.

Conceptualizing self-efficacy

In general, self-efficacy refers to how confident an individual feels about handling particular tasks, challenges, and contexts (Bandura, 1997). It is widely considered to be derived from Bandura's (1986) Social Cognitive Theory (SCT). Bandura (1994) defines self-efficacy as people's beliefs "about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (p.71). It is generally reported that individuals with higher self-efficacy perceive difficult tasks as meaningful challenges, despite the fact that others may find similar tasks discouraging. In Bandura's (1994) understanding, high self-efficacy "fosters intrinsic interest and deep engrossment in activities" (p.71); on the contrary, a lack of self-efficacy may cause people to have low aspirations, slacken their efforts, and give up easily. In addition, some researchers (e.g., Girasoli & Hannafin, 2008) have further indicated that learners' cognitive processes can be influenced by self-efficacy.

Furthermore, as proposed by Pajares and Schunk (2001), instead of being evaluated in general, research regarding self-efficacy should be assessed at a domain-specific or task-specific level because such measures may have greater validity and predictive relevance. In other words, domain-specific self-efficacy assessment, such as asking students to state their confidence in learning mathematics or writing, is more explanatory and predictive than omnibus measures and preferable for making general academic judgments (Pajares, 1996).

Levels of self-efficacy are usually considered to have strong validity for specific task domains, and most of the findings have suggested that self-efficacy is positively related to learners' performance. That is, a strong sense of self-efficacy can enrich human achievement in many ways (Karsten & Roth, 1998). For example, Caprara et al. (2008) indicated that the lower the decline in self-efficacy, the higher the grades and the greater the likelihood of remaining in high schools. Hoffman and Spathariu (2008) similarly demonstrated the positive effects of self-efficacy on problem-solving efficiency. Based on the above literature, when exploring the relationship between self-efficacy and IBL, it becomes important to interpret self-efficacy carefully from different perspectives.

Self-efficacy in IBL environment

Recently, a great amount of research relating to self-efficacy has been carried out in the educational research community. Nevertheless, different researchers have observed learners' self-efficacy from a variety of perspectives. As a result, prior to describing relevant applications of self-efficacy in the IBL research, it may be helpful to identify appropriate definitions for different types of self-efficacy. In general, academic self-efficacy (ASE) pertains to a student's perception of academic learning (Girasoli & Hannafin, 2008), while computer self-efficacy (CSE) is defined as an individual's perceived confidence regarding his/her ability to use a computer (Compeau & Higgins, 1995; Murphy, Coover, & Owen, 1989). Likewise, general Internet self-efficacy describes people's perceptions about their own abilities to use the Internet (Tsai & Tsai, 2003), whereas IBL self-efficacy represents individuals'

confidence and self-belief in their ability to master an online course or online learning activity (Yukselturk & Bulut, 2007). For the purpose of this review and for the consistency of terminology, learners' general Internet self-efficacy is named as their ISE (Internet self-efficacy) in the present study.

In comparison with the development of computers, Internet technology is viewed as a relatively innovative invention. Therefore, before attempting to interpret the conceivable relations between self-efficacy and IBL, it is meaningful to discuss the relevant findings concerning CSE. Marakas Yi, and Johnson (1998) defined CSE as "an individual's perception of efficacy in performing specific computer related tasks within the general computing domain" (p. 127). Thus, CSE can be considered a domain specific measure of self-efficacy that reflects a person's belief in his/her ability to perform specific computer tasks. Consistent with the original concept of the self-efficacy theory, CSE is developed over time and is thought to have influences on the consequence of learners' interactions with computers when facing obstacles (Compeau & Higgins, 1995; Murphy, Coover, & Owen, 1989). Compared with ISE-related work, more extensive literature on CSE has been published. For instance, Moos and Azevedo (2009) conducted a comprehensive literature review on the relations between computer-based learning environments and CSE. Based on their report, a number of scholars have distinguished CSE into disparate dimensions. For example, Marakas, Yi, and Johnson (1998) and Marakas, Johnson, and Clay (2007) have divided CSE into two distinct levels: general computer self-efficacy, which assesses learners' general beliefs about their computing skills (e.g., their confidence in using software to complete a computing job), and application-specific self-efficacy, which assesses confidence in using specific applications (i.e. confidence in the ability to rename a file in specific applications such as Excel or Word).

Moos and Azevedo's (2009) work synthesized the studies including those empirically examined factors related to CSE and the relationship among CSE, learners' learning outcomes, and learning processes in the computer-based learning environment. On the basis of their findings, both behavioral and psychological factors were found to be positively related to CSE, which is related to students' learning outcomes in computer-based learning environments. Besides, it was shown that this relationship may change with students' acquisition of skills or knowledge. Finally, users' CSE might be related to their navigational paths in computerized learning environments. Different from Moos and Azevedo's review on CSE, the present research specifically focuses on those studies pertaining to the relations between self-efficacy and IBL environments.

Method

Paper selection

In this study, the Social Science Citation Index (SSCI) database from 1999 to 2009 was used for paper selection using the following keywords for topics: Internet AND self-efficacy; web AND self-efficacy; network AND self-efficacy; e-learning AND self-efficacy; online AND self-efficacy. The first phase of the search produced 489 articles. Studies published from 1999 to 2009 were selected because Internet technology is considered to have been widely implemented in the educational realm since 1999. To illustrate, in 1999, the UNESCO Institute on Information Technologies in Education initiated and began the project the Internet in Education (UNESCO, 2003). Moreover, IBL was defined as those learning activities taking place in an Internet-based setting. Then, the data gathering procedure was directed to the subsequent selection derived from the criteria determined by three experts in the field of educational technology. The selection criteria were comprised of three principles: (a) the major purpose of the study must include at least one component probing the role of self-efficacy in any kind of IBL condition, (b) the study design should be based on an empirical methodology, and (c) the main findings of the research must be related to learning and must elaborate the application of self-efficacy in an Internet-based setting. Abstracts were first reviewed and articles were then limited according to these principles. Then, full papers were examined for the relevancy to this review. On the basis of the previously mentioned criteria and three rounds of expert panel discussions for the validation of the selection, 46 articles remained for the current review.

Review framework

Four educational researchers examined the 46 papers selected, conducted content analyses by summarizing the major findings of the studies, and after two rounds of discussions, concluded three categories for this review, which could cover almost all of the topics under investigation. The first category consisted of the studies relating to learners'

general Internet self-efficacy (ISE); the second category included the investigation exploring the interplay between learners' academic self-efficacy and the Internet-based learning (ASE&IBL). The third category contained research probing learners' IBL self-efficacy (IBLSE), that is, learners' self confidence in their participation and their expected performance in an IBL setting.

Table 1. The review framework for the research regarding self-efficacy and Internet-based learning

| Category | Subcategory 1 Role of self-efficacy in IBL | Subcategory 2 Alteration of self-efficacy in IBL |
|---|--|---|
| ISE (Internet Self-Efficacy) | To investigate relations between learners' ISE and learning processes or outcomes in IBL | To probe how learners' ISE may be altered in IBL |
| ASE&IBL (Academic Self-Efficacy and Internet-Based Learning) | To investigate the interplay between learners' ASE and learning processes or outcomes in IBL | To probe how learners' ASE may be altered in IBL |
| IBLSE (Internet-Based Learning Self-Efficacy) | To investigate relations between learners' IBLSE and learning processes or outcomes in IBL | To probe how learners' IBLSE may be altered in IBL |

Moreover, similar to the framework applied in other reviews (e.g. Lee *et al.*, in press; Tallent-Runnels *et al.*, 2006), two subcategories were further drawn. The first subcategory comprised the studies exploring the relationship between students' self-efficacy and their learning process or learning outcomes in IBL conditions. Meanwhile, the second subcategory was made up of the research probing how students' self-efficacy might be altered among different IBL contexts. Table 1 provides an outline of the research framework of the present study.

Review Results

Internet Self-efficacy (ISE)

According to Table 1, studies in the ISE category were divided into two subcategories. On the one hand, the investigation between learners' ISE and their learning processes or outcomes in the IBL condition was explored. On the other hand, an amount of research was utilized to probe how learners' ISE may have altered in IBL settings. A complete list of research involving learners' ISE is summarized in Table 2.

Table 2. Summary of ISE research in alphabetic order. (* = included in both ISE and ASE&IBL categories; exp=experimental design)

| Author | Purpose | Participants | Method |
|-----------------------------|---|--|---------------|
| *Brown <i>et al.</i> (2003) | Find ISE changes through GlobalEd Project | 234 high school students in US | survey |
| Chiou & Wan (2007) | To investigate changes of ISE on information searching in the Internet-based condition | 136 college students in Taiwan | exp |
| Chu & Tsai (2009) | To build a model explaining ISE's influence on adult learners' preferences for IBL | 541 adult learners in Taiwan | survey |
| Hong (2006) | To assess ISE's role in health-related online search | 84 US university student | exp |
| *Joo <i>et al.</i> (2000) | To test the applicability of self-efficacy theory to the context of IBL | 152 junior high school students in Korea | survey |
| Lam & Lee (2006) | To investigate the role of ISE and outcome expectations in older adults' usage of Internet | 1000 adults in Hong Kong | exp |
| Liang & Tsai (2008) | To explore the relations between ISE and preferences for constructivist IBL | 365 college students in Taiwan | survey |
| Lu <i>et al.</i> (2007) | To explore ISE's effect on students' likelihood of using Internet-based systems to seek information | 229 university students in US | survey |
| O'Malley & Kelleher(2002) | To examine PR students' ISE in either geographically dispersed or local teams | 55 university students in US | quasi-exp |

| | | | |
|--------------------------------|---|------------------------------------|------------|
| Peng <i>et al.</i> (2006) | To explore ISE & perceptions of Internet | 1417, university, Taiwan | survey |
| Schmidt & Ford (2003) | To evaluate the impact of meta-cognition activities on ISE in IBL | 79 undergraduates in US | exp |
| *Thompson <i>et al.</i> (2002) | To examine the relationship between learners' self-efficacy and their search task performance | 90 undergraduates in US | exp |
| Torkzadeh & Van Dyke (2001) | To develop an appropriate ISE instrument | 277 undergraduates in US | survey |
| Torkzadeh & Van Dyke (2002) | To examine the relationship among users' training, attitude, and ISE | 189 university students in US | survey |
| Torkzadeh <i>et al.</i> (2006) | To develop and examine a contingency model of ISE | 347 university students in US | survey |
| Tsai & Tsai (2003) | To examine ISE's effect on information search strategies for IBL science learning | 8 university freshmen in Taiwan | case study |
| Wu <i>et al.</i> (2006) | To investigate the relationship between students' attitude and their ISE | 1313 university students in Taiwan | survey |
| Yang <i>et al.</i> (2007) | Explore how ISE mediates IAN & IBL intention | 368, university, Taiwan | survey |
| Yi & Hwang (2003) | To predict the use of IBL by incorporating ISE | 109, university, US | survey |

Relations between ISE and learning processes/outcomes

It was found that some studies had been conducted to assess learners' basic perceptions of ISE in IBL prior to the investigation of the relationship between learners' ISE and their learning processes or outcomes in the IBL condition. For instance, Torkzadeh and van Dyke (2001), Wu and Tsai (2006), and Peng, Tsai, and Wu (2006) developed a number of questionnaires in order to assess learners' basic perceptions of ISE. Torkzadeh and van Dyke (2001) used 277 responses from university students to develop and validate a 17-item ISE scale. Statistical analysis supported a three-factor model, including *surfing/browsing*, *encryption/decryption*, and *system manipulation*. According to their report, the first factor assessed learners' confidence in surfing, browsing or finding information in an IBL setting; the second factor assessed learners' confidence in decrypting or encrypting messages in an Internet-based setting; the third factor assessed learners' confidence in operating an IBL system. Evidence of reliability and construct validity were indicated in their study.

Similarly, to find out learners' fundamental perceptions of ISE, Tsai and his colleagues (Peng, Tsai, & Wu, 2006; Wu & Tsai, 2006) divided ISE into two types: *general Internet self-efficacy* and *communicative Internet self-efficacy*. General self-efficacy addresses students' Internet self-efficacy for basic functions or purposes (e.g. *I can search for information on the Internet by using keywords*), whereas communicative self-efficacy probes their efficacy for Internet-based communication or interaction (e.g. *I think I can talk to others in online chat rooms*). With a sample of 1,313 university students in Taiwan, Wu and Tsai (2006) found that students' Internet attitudes are highly correlated with not only general ISE but also with communicative ISE. It was suggested that students' Internet attitudes could be viewed as one of the important indicators for predicting ISE. In a similar way, Peng, Tsai, and Wu (2006) investigated 1,417 Taiwanese university students and proposed that students perceiving the Internet as a leisure tool (e.g. as a tour or a toy) showed higher communicative ISE than those using the Internet simply as functional technology.

In addition to probing learners' basic perceptions of ISE, Tsai and his research team also attempted to identify the conceivable relationship between learners' ISE and their learning processes or learning outcomes (Chu & Tsai, 2009; Liang & Tsai, 2008; Tsai & Tsai, 2003). For example, to examine the role of students' ISE in their information searching strategies in an IBL setting, Tsai and Tsai (2003) conducted 8 in-depth case studies and concluded that high ISE students had better information searching strategies and learned better than those with low ISE in the Internet-based condition. Moreover, in an attempt to explore the relationship between learners' ISE and their preferences for IBL, Liang and Tsai (2008) surveyed 365 Taiwanese college students and revealed that students with higher ISE (e.g. "I can search for information on the Internet by using keywords.") showed greater preferences for IBL which they could use with ease; however, students with higher communicative ISE (e.g. "I think I can talk to others in online chat rooms.") tended to display relatively weaker preferences for inquiry learning in IBL.

Recently, to establish a theoretical model to explain factors that might influence adult learners' preferences for constructivist IBL settings, Chu and Tsai (2009) gathered data from 541 participants enrolled in adult education institutes for structural equation modeling (SEM) analyses. The results revealed that ISE plays a mediating role in the relationships between Internet usage and the participants' preference for IBL, and indicates that with augmented time spent on Internet practice, adult learners' ISE, strengthening their preferences for IBL, may be increased.

Besides exploring how ISE might be associated with learner's learning processes, a certain amount of research has investigated the relationship between ISE and the subsequent use of Internet-based systems. For instance, Yi and Hwang (2003) extended the technology acceptance model by incorporating ISE to predict the use of the Blackboard system by surveying 109 university students, and concluded that ISE positively influences the decision to use Internet-based technology and subsequent actual use. Contradictory to Yi and Hwang's positive results, in an attempt to explore factors influencing students' likelihood of using the Internet to seek information, Lu *et al.* (2007) found that respondents' ISE had no significant association with their intentions to seek information on the Internet after surveying 229 international students. Similarly, with a sample of 368 undergraduates, Yang *et al.* (2007) found that the anxiety of Internet use negatively influenced ISE, whereas ISE did not significantly affect the intention to use Internet sites. Lam and Lee (2006) inquired into the role of ISE and outcome expectations in learners' usage of the Internet by a longitudinal study among 1,000 seniors in Hong Kong. Their findings generally validated the effects of ISE and outcome expectations on the Internet usage intention.

Finally, Hong (2006) explored the effects of ISE and search task specificity on the outcomes and task perseverance of finding online health-related sites that contained attributes of website accountability. In the study, 84 US university students conducted two search tasks (general and specific) that varied in the degree of task difficulty. The results showed that high ISE participants located sites higher in website accountability in the general search task than their low ISE counterparts. Besides, the participants with high ISE demonstrated more task perseverance than those with low ISE.

Alteration of ISE

A number of studies were found to utilize experimental designs and to examine how learners' ISE might be changed through an intervention or training. Take Schmidt and Ford's (2003) study for example, 42 undergraduate students received a brief introduction to metacognitive practices, in which trainees were informed to more frequently and accurately reflect on what they were learning through the program before creating Web pages, while 37 participants in the control condition began the Web-page creation training immediately. Consistent with their expectations, learners reporting greater levels of metacognitive activity during training had higher levels of ISE when compared with their counterparts.

Similarly, O'Malley and Kelleher (2002) required 55 university students majoring in public relations to develop a statement and measured their ISE before, immediately after, and 7 weeks after working in either geographically dispersed (Kansas and Hawaii) or local (Kansas only) teams. In the experimental section, two participants from Kansas State and two from Hawaii were randomly assigned to each group. The participants in the control section were also randomly assigned, but they were not assigned to collaborate with students from Hawaii. The results revealed that learners' ISE increased over time regardless of the experimental conditions.

Torkzadeh and his team members conducted several experimental studies to probe how ISE might be elevated or related to learners' Internet attitudes. Torkzadeh and van Dyke (2002) as well as Torkzadeh, Chang, and Demirhan (2006) reported the effects of training on learners' ISE and their computer user attitudes. With a 17-item ISE scale developed and validated in 2001, Torkzadeh and van Dyke (2002) reported on the effects of training on students' ISE and their computer user attitudes by utilizing questionnaires with a sample of 189 university students. Training was considered an important way of improving computer-related self-efficacy (Compeau & Higgins, 1995; Marakas, Yi, & Johnson, 1998). Therefore, the study collected questionnaire responses at both the beginning and end of an introductory computer training course. The content of the training program was not clearly reported in Torkzadeh and van Dyke's (2002) study, but they suggested that the training significantly improved the students' ISE. Besides, respondents with a favorable attitude toward computers were found to have higher ISE scores than those with an unfavorable attitude; and male respondents consistently reported higher than females for ISE on both the pre- and post-training scores.

Torkzadeh, Chang, and Demirhan (2006) developed and examined a contingency model of learners' CSE and ISE. With measures of user attitude, computer anxiety, computer self-efficacy, and Internet self-efficacy, the authors analyzed the survey responses of 347 university students from multiple sections of a training course like information technology infrastructure and decision support systems.. The result suggested that the training programs significantly improved learners' CSE and their ISE. Besides, respondents with favorable attitudes toward computers improved their ISE significantly more than those with unfavorable attitudes. Respondents with low computer anxiety improved both their CSE and ISE significantly more than those with high computer anxiety; however, the interaction effect between attitude and anxiety was only significant for the CSE scores but not for the ISE scores.

Finally, Chiou and Wan (2007) investigated the change of ISE on information searching in an Internet-based condition. The students receiving low-difficulty manipulation (i.e., allowing a longer search period) obtained a higher level of ISE, whereas those receiving high-difficulty manipulation (i.e., allowing a shorter search period) possessed a lower level of ISE. The results indicated that the enhancement effect of positive task experience (such as low-difficulty tasks) on self-efficacy was more pronounced for individuals with lower levels of ISE on information searching in Internet-based settings, whereas the deteriorating effect of negative experience was more prominent for individuals with higher levels of ISE on information searching in the Internet-based condition.

Summary of ISE research

In conclusion, research found in the ISE category mainly focused on developing and validating methods of assessing learners' ISE and exploring its relationship with those factors likely to play a role in students' learning processes or outcomes in the IBL condition. To name a few, the relations among students' ISE and their attitudes, strategies, and preferences were examined.

Besides, some researchers also paid attention to gender-related issues while making efforts to link ISE with the abovementioned constructs. It is generally believed that computer-related tasks are more advantageous for males than females (Li & Kirkup, 2007). Although some studies (e.g. Wu & Tsai, 2006; Torkzadeh & van Dyke, 2002) have actually found that male students reveal better ISE than their female counterparts, a report of 234 high school participants conducted by Brown et al. (2003) suggested that either boys' or girls' ISE regarding a specific simulation ILE named GlobalEd Project had revealed similar patterns.

Interestingly, when taking measures to explore learners' levels of ISE, a majority of the studies (e.g. Joo, Bong, & Choi, 2000; Thompson, Meriac, & Cope, 2002; Tsai & Tsai, 2003) adopted the behavior or the performance of search tasks as students' learning outcomes. To illustrate, Thompson, Meriac, and Cope (2002) found positive correlation between ISE and the number of correct search results produced. Joo, Bong, and Choi (2000) stated that students' scores on the Internet-based search tests were significantly and positively predicted by their ISE. Hence, it may be suggested that search tasks is regarded as the most commonly implemented IBL activities at the present stage.

Interplay between academic self-efficacy and Internet-based learning (ASE&IBL)

According to Bandura (1997), ASE is defined as students' expectations of how successful they will be in the classroom. There is no doubt that issues relating to ASE have been extensively researched (Pajares, 1996; Schunk & Pajares, 2002); however, most of the investigations are not related to the learning occurring in Internet-based contexts. Therefore, in the present review, the ASE&IBL category was specifically drawn to discuss the interplay between learners' ASE and IBL settings.

At first, it was found that a variety of methods and objectives characterized the research comprising the ASE&IBL category. To name a few, ASE was disclosed to be associated with goal orientation (Sins *et al.*, 2008), self-regulated learning (Crippen & Earl, 2007; Yukselturk & Bulut, 2007; Joo *et al.*, 2000), and motivational beliefs (Yukselturk & Bulut, 2007; Tai, 2006) in IBL settings. Moreover, it was found that, on the one hand, most of the non-experimental studies have investigated the relationship between learners' ASE and their motivational constructs influencing pupils' IBL processes or outcomes. On the other hand, the experimental research was inclined to the comparisons among diverse types of learning environments. More specifically, to probe potential variations, some were conducted

within a variety of forms of the IBL setting, but some were implemented within both Internet-based and traditional F2F learning conditions. It was proposed that different extents of students' ASE may have been derived from the differences in these learning environments. The studies involving ASE&IBL issues are summarized in Table 3.

Table 3. Summary of ASE&IBL research in alphabetic order. (* = included in both ISE and ASE&IBL categories; exp=experimental design)

| Author | Purpose | Participants | Method |
|----------------------------------|---|--|------------------|
| *Brown <i>et al.</i> (2003) | To find ISE changes through GlobalEd Project | 234 high school students in US | survey |
| Crippen & Earl (2007) | To create an Internet-based program which improves user performance and supports well-structured problem solving | 66 university students in US | quasi-exp |
| Farel <i>et al.</i> (2001) | To show IBL effect on professional development and practice | 28 staffs in US | exp |
| Francescato <i>et al.</i> (2006) | To compare learners' ASE between collaborative F2F and IBL groups | 50 university students in Italy | exp |
| Francescato <i>et al.</i> (2007) | To compare learners' self-efficacy in developing professional skills in collaborative F2F and IBL courses | 166 university students in Italy | survey interview |
| *Joo <i>et al.</i> (2000) | To test the applicability of self-efficacy theory to contexts of IBL | 152 junior high school students in Korea | survey |
| Kitsantas & Chow (2007) | To examine how ASE varied across four different instructional environments | 472 college students in US | quasi-exp |
| Meyer <i>et al.</i> (2002) | To assess the impact of using the Internet-based setting where elders provided tutoring for students | 12 adults (62-80 yrs old) and 60 5th-graders in US | exp |
| Sins <i>et al.</i> (2008) | To test relations among learners' AS, goal orientation, cognitive processing, and achievement in collaborative IBL | 60 pre-university science majors in Holland | survey |
| Tai (2006) | To test how the effect of training framing from supervisors on trainees' AS may influence the overall effectiveness | 126 employees in Taiwan | exp |
| *Thompson <i>et al.</i> (2002) | To examine the relationship between learners' self-efficacy and their search task performance | 90 undergraduate students in US | exp |
| Waldman (2003) | To examine the role AS played in students' use of the library's electronic resources | 340 university students in US | survey |
| Yukselturk & Bulut (2007) | To examine relationship among learner selected variables, AS, self-regulated learning, and their success in IBL | 80 online students in Turkey | survey interview |

Relations between ASE and IBL processes/outcomes

As stated, a number of studies (i.e. Sins et al., 2008; Waldman, 2003; Yukselturk & Bulut, 2007) have explored how ASE, coupled with other motivational constructs, plays a role in students' successful IBL. That is, they have sought to find out how ASE may be associated with students' IBL processes or IBL outcomes.

Sins et al. (2008) tested 60 11th-grade students' conceptual model of the relationship between students' achievement goal orientation and their ASE with respect to a modeling task in an Internet-based setting, where learners could collaborate online by means of a synchronous chat on inquiry assignments for science courses. The study found that learners' mastery-approach goal orientation and their ASE were both positively related to their achievement in the modeling task and their use of deep cognitive processes.

Similarly, in an attempt to find out what might influence students' IBL of programming, Yukselturk and Bulut (2007) analyzed and examined the relationship among 80 online learners' selected variables (i.e. gender, age,

educational level, locus of control, and learning style), motivational beliefs (i.e. intrinsic goal orientation, extrinsic goal orientation, control beliefs, task values, self-efficacy, and test anxiety), self-regulated learning components, and their success in an Internet-based setting. The study result showed that although learners' ASE and their intrinsic goal orientation beliefs were correlated with their IBL success, they did not enter the final prediction model in the regression analyses. This finding was somewhat contradictory to Pintrich and de Groot's (1990) research outcomes, in which respondents' ASE and their intrinsic motivation significantly affected their achievement.

In a similar way, to encourage students to use library facilities and electronic resources, Waldman (2003) conducted a study to understand what factors may have promoted students to seek out information in a library setting. Previous research showed that students' ASE might be related to their academic achievement outcomes (Bandura, 1997; Pajares, 1996). Therefore, Waldman (2003) analyzed 340 university freshmen's responses concerning their library/computer usage and their ASE. The research finding showed that the students who expressed interest in learning about the library's electronic resources were more likely to have higher ASE for completing the learning task. Moreover, students with higher ASE for completing the task tended to use the library more often. The study outcome was consistent with other research on self-efficacy (e.g. Ren, 2000), suggesting that self-efficacious students have a tendency to be more active in academic work and to use resources available to them.

Finally, Thompson, Meriac, and Cope (2002) examined the relationship between learners' self-efficacy (including both general ISE and ASE) and their search task performance in the IBL condition. A total of 90 participants were required to search the Internet and to list the names of the industrial-organizational psychologists they found. The findings indicated that the improvement in both ISE and ASE could lead to higher online performance.

Alteration of ASE by IBL

Whereas the research discussed above mainly deals with how ASE may influence or be related to students' learning processes and outcomes in an Internet-based setting, the following studies observe how ASE may be altered in different Internet-related learning conditions. It was found that some studies measured the learners' ASE within IBL conditions only (e.g. Crippen & Earl, 2007; Meyer et al., 2002), while others intended to compare learners' differences of ASE between traditional F2F classroom settings and Internet-based situations (e.g. Francescato et al., 2006, 2007; Kitsantas & Chow, 2007).

Crippen and Earl (2007), Meyer et al. (2002), and Tai (2006) have evaluated participants' ASE among Internet-based conditions only. Crippen and Earl (2007) described a quasi-experimental study, wherein expert modeling was believed to improve ASE, and worked examples served as expert models in their study. A total of 66 students were randomly assigned to one of three conditions: a worked example group, a worked example/self-explanation group, and the control group. In the end, the combination of worked example with self-explanation prompt was reported to improve students' performance, problem solving skills, and ASE in terms of whether personal goals were achieved.

Meyer et al. (2002) assessed the impact of using a structured strategy as a base for an intergenerational Internet tutoring program, in which 12 older adults provided tutoring for 5th-grade students to learn the strategy through an instructional Internet-based system. The structured strategy was considered to allow readers to build mental representations similar to the text's hierarchical organization of important ideas. Sixty students were randomly assigned to one of three groups: (a) a tutoring group, in which the students worked on the Internet-based system using the structured strategy with a tutor; (b) a group in which the students worked independently on the same Internet-based instruction without a tutor; and (c) a control group, in which the students did not receive instruction in the structured strategy. The results showed that both tutors and children in the structured strategy group with tutors increased their ASE.

Likewise, to examine the effects of training framing from supervisors on trainees' ASE and training motivation in IBL, Tai (2006) surveyed 126 employees entering a training program introducing computer software operation and design, and further tested how these variables may have influenced the overall training effectiveness. The 126 employees were asked to complete a series of questionnaires at the beginning, the midpoint, and the end of the course. The results indicated that supervisors training framing could be used to predict trainees' ASE, which subsequently affected their reactions, learning, and motivation.

As mentioned earlier, other researchers have intended to identify the differences in learners' ASE between traditional F2F classrooms and IBL settings. The exploration of such differences may illustrate the role or the effects of IBL on students' ASE. For example, Francescato *et al* (2006, 2007) compared learners' self-efficacy between traditional F2F and IBL conditions within a computer-supported collaborative learning (CSCL) setting in particular. Francescato *et al.* (2006) implemented a pilot study, in which 50 psychology major students were required to learn the same material in the F2F and Internet-based classroom settings. The results indicated that participants in both groups achieved similar growth in their levels of ASE, social self-efficacy, and self efficacy for problem solving. Collecting data from a different sample, Francescato *et al.* (2007) conducted the other study with 166 students in similar experimental conditions. Different from the previous study, the results of the second research found statistically significant increases only in learners' social self-efficacy and self-efficacy for problem solving for both groups, but not in their ASE. On the basis of their study outcomes, no significant increase in learners' ASE for both F2F and Internet-based CSCL conditions was found; thus, it may be suggested that the Internet-based CSCL environment could be regarded as efficient as traditional F2F classroom settings in increasing learners' social self-efficacy and self-efficacy for problem solving, but the effects on ASE may vary across the two studies.

Kitsantas and Chow (2007) examined how college students' help-seeking behavior varied across three different instructional learning environments. A total of 472 students enrolled in distance, distributed, and traditional classes were queried about their help-seeking preferences, help-seeking tendencies, personal threat in seeking help, and ASE. The research findings showed that, regardless of the class in which they were enrolled, the students' academic achievement was positively associated with their ASE. These results were consistent with previous research findings, in which ASE was positively inter-correlated and predicted achievement, and students with higher ASE for successful problem solving displayed greater performance monitoring and persisted longer than those with lower ASE (Pinrich & de Groot, 1990).

Finally, in an attempt to examine the changes in self-efficacy through Internet-based courses, one study concerning medical professional training was reported. Farel, Umble, and Polhamus (2001) discussed the effect of an analytical skills training course on medical professional development and practice. Through a one-year Internet-based program, the study found that 28 participants' ASE increased significantly, suggesting that the Internet-based analytical and technical training initiatives could offer a promising means for reaching public health professionals, and provide an alternative opportunity for off-site workshops. With the IBL, in-service professional practitioners could acquire easier access to and adoption of training to meet their needs, which may have led to greater motivation as well as increased ASE.

Summary of ASE&IBL research

Because students' perceptions of ASE are considered to be important in their use of self-regulated strategies (Zimmerman & Martinez-Pons, 1990), an amount of research (e.g. Crippen & Earl, 2007; Yukselturk & Bulut, 2007; Joo, Bong, & Choi, 2000) has probed the potential interplay between learners' ASE and their self-regulated learning activities in Internet-based settings. Furthermore, a number of studies (e.g. Joo, Bong, & Choi, 2000; Brown, et al., 2003; Thompson, Meriac, & Cope, 2002) were found to query not only learners' ASE but also their ISE. For instance, Joo, Bong, and Choi (2000) tested the applicability of self-efficacy theory to the contexts of a specific Internet-based condition, in which learners had to conduct several search tasks. They found that learners' perceptions of ASE could predict their performance measured by written tests, whereas their perceptions of ISE were significant in predicting their search test performance. These results give evidence to support that learners' ASE is more associated with achievement measured by a conventional assessment mode, while learners' ISE is related to their performance in operating Internet-related functions. Similarly, Brown et al. (2003) discussed gender issues in terms of learners' ASE and ISE. They found no difference between the two genders and concluded that both boys and girls revealed similar patterns of responses for both kinds of self-efficacy. Finally, Thompson, Meriac, and Cope (2002) examined learners' ASE and ISE in the IBL condition and suggested that improvement in both self-efficacies could lead to higher online performance. This conclusion was consistent with Bandura's (1986) assertion, indicating a reciprocal interaction between learners' self-efficacy and their performance. Therefore, when Girasoli and Hannafin (2008) reviewed the potential importance of designing scaffolds in the Internet-based condition, they suggested that both students' ASE and their general ISE should be taken into consideration and intentionally promoted.

Internet-based learning self-efficacy (IBLSE)

The category of IBLSE is made up of the research which examines learners' confidence in their participation and their expected consequent performance particularly derived from IBL activities. A number of features were found among the studies.

First of all, rather than developing a full instrument particularly evaluating learners' IBLSE, most researchers only included a factor with a limited number of questions in their surveys. Several reasons may account for a lack of relevant research. First, a large number of participants are required for the development of surveys, and it may somehow be difficult for researchers to have access to a large group of individuals possessing complete IBL experiences. IBLSE-related instruments can be context dependent and susceptible to a specific kind of IBL programs. Thus, it becomes difficult for other researchers to validate a questionnaire previously utilized in other research with different types of IBL activities. Thus, IBLSE-related instrument or factor may only play a minor role in many of the studies (e.g. Chang & Tung, 2008; Liaw, 2008; Tsai, 2009).

Besides, compared with the research concerning ISE or ASE&IBL, studies relating to IBLSE seem to be widely applied to certain established models. A certain amount of research was found to comply with the Technology Acceptance Model (TAM) and its associated implementation regarding the IBLSE construct.

Finally, very little research aimed to investigate the changes of learners' IBLSE. Instead of implementing experimental designs, a large amount of research has merely presented relational/co-relational data within this category (e.g. Artino, 2008; Johnson, Hornik, & Salas, 2008; Wang & Newlin, 2002). A limited number of experimental studies may suggest that, rather than observing the possible differences among diverse learning conditions, IBLSE-related research seems to have explored the interplay between learners' self-efficacy in terms of IBL and their academic outcomes in such learning activities, or their satisfaction with Internet-based programs. A complete list of studies involving IBLSE issues is provided in Table 4.

Table 4. Summary of IBLSE research in alphabetic order. (exp=experimental design)

| Author | Purpose | Participants | Method |
|------------------------------|--|---|-----------|
| Artino (2008) | To investigate the relations between learners' IBLSE and their satisfaction | 646 undergraduates in US | survey |
| Bates & Khasawneh (2007) | To propose a mediated model where a set of antecedent variables influenced students' IBLSE | 288 university students in US | survey |
| Bolman <i>et al.</i> (2007) | To investigate the usability of the IBL system | 808 unders, friends & families in Holland | exp |
| Chang & Tung (2008) | To add IBLSE as one of the factors to propose a new hybrid TAM. | 212 undergraduates in Taiwan | survey |
| Choi <i>et al.</i> (2007) | To suggest an IBL success model based on flow theory | 223 vocational school students in Korea | quasi-exp |
| Johnson <i>et al.</i> (2008) | To develop a model by adding social presence to learners' IBLSE. | 345 university students in US | survey |
| Lee (2006) | To investigate factors affecting the adoption of the IBL through TAM | 1085 university students in Taiwan | survey |
| Lee & Lee (2008) | Suggest a research model based on relations of IBLSE and IBL system perception | 225 unders in Korea | survey |
| Liaw <i>et al.</i> (2007). | To explore instructors' and learners' attitudes toward IBL | 30 instructors & 168 unders in Taiwan | survey |
| Liaw (2008) | To examine relations among learner satisfaction, IBLSE, and IBL effectiveness | 424 university students in Taiwan | survey |
| Moneta <i>et al.</i> (2007) | To evaluate students' affective learning in IBL | 414 undergraduates in Hong Kong | quasi-exp |
| Ong & Lai (2006). | To explore gender differences among dominants affecting IBL acceptance | 156 employees in Taiwan | survey |
| Park (2009) | To investigate students' adoption of IBL via SEM technique with LISREL program | 628 university students in Korea | survey |

| | | | |
|----------------------|---|-----------------------------------|-----------|
| Tsai (2009) | To examine differences between conceptions of learning and of IBL | 83 college students in Taiwan | interview |
| Tung & Chang (2008a) | To study nursing students' behavioral intentions to use IBL | 267 university students in Taiwan | survey |
| Tung & Chang (2008b) | To propose a new hybrid technology acceptance model (TAM) | 228 university students in Taiwan | survey |
| Wang & Newlin (2002) | To investigate if students' IBLSE would predict their performance | 122 college students in US | quasi-exp |

Relations between IBLSE and IBL processes/outcomes

As stated, most of the findings revealed in the IBLSE-related studies pertain to certain established models. Consequently, IBLSE is frequently utilized as a predicting factor which may influence or be related to students' learning processes or outcomes. Moreover, it was also noted that a majority of the research pays attention to the application of the Technology Acceptance Model (TAM). For example, Park (2009) surveyed 628 university students' adoption of IBL using the structural equation modeling technique. A general structural model including IBLSE was developed, and IBLSE was found to be the most important construct for the participants' intention to use e-learning. Similarly, Lee (2006) probed the factors affecting the adoption of an IBL system by surveying 1,085 Taiwanese university students online. The research evidence was found to lend support to the original TAM findings, in which students' IBLSE was demonstrated to significantly relate to their perceived ease of use. Hence, it was considered important to develop an easy-to-use system and to increase participants' IBLSE.

In a similar way, on the basis of questionnaire responses collected from 67 female and 89 male employees in six international companies based in Taiwan, Ong and Lai (2006) explored gender differences in perceptions and relationships among factors affecting the acceptance of IBL. The research findings showed that although females' rating of IBLSE was lower than males', females' perception of IBLSE played a more important determinant role in affecting their behavioral intention to take part in IBL. It was hence suggested that gender issues should be taken into consideration when developing relevant theories.

Furthermore, in a series of studies probing learners' intention to engage in IBL, Chang and Tung (2008) and Tung and Chang (2008a, 2008b) proposed new hybrid models in which TAM was combined with the innovation diffusion theory. According to their research results, IBLSE was one of the critical factors which may have an impact upon students' behavioral intentions. At first, Chang and Tung (2008) combined the innovation diffusion theory and TAM, and added two research variables comprising of perceived system quality and IBLSE to study students' behavioral intentions to use an IBL course. With an analysis of questionnaire responses from 212 undergraduate students who were using online learning course websites in Taiwan, the study found that IBLSE had a positive effect on students' behavioral intention to use the online learning course websites.

Similarly, derived from the questionnaire responses from 267 nursing students of six universities in Taiwan, Tung and Chang (2008a) reported that IBLSE had a positive effect on learners' behavioral intention to use the Internet-based nursing program. Finally, Tung and Chang (2008b) added four variables, including computer anxiety, IBLSE, perceived financial cost, and perceived information quality. Based on 228 questionnaires collected from nursing students who had taken Internet-based courses in Taiwan, they found the more confident students were in their ability to use IBL (i.e., higher IBLSE), the more likely they were to take part in Internet-based courses.

Likewise, deriving from the notion of TAM, a three-tiered Technology Use Model (3-TUM) was developed by Liaw, Huang, and Chen (2007) and Liaw (2008) to investigate how individuals' IBLSE influences their satisfaction and behavioral intention regarding IBL programs. In their studies, 3-TUM was defined as integrated multidisciplinary perspectives comprising motivation, social cognitive theory, theory of planned behavior, and TAM. In Liaw, Huang, and Chen's (2007) study, a total of 30 instructors were asked to answer a series of questionnaires. The result indicated that instructors' behavioral intention to use the IBL program was positively influenced by their perceived IBLSE. Similarly, derived from 424 university students' survey responses, Liaw (2008) stated that users' perceived IBLSE played a positive role in determining students' satisfaction with and behavioral intention to use Internet-based systems.

Though Lee and Lee (2008) did not directly refer their study to the notion of TAM, they proposed a research model adopting IBLSE as a moderating variable to investigate learners' perceptions of the quality of an IBL system. Their study result indicated that higher IBLSE group was more sensitive to the effectiveness and usefulness of the system and was aware of the contextual information quality (e.g. the variety of the lectures) than those with lower IBLSE. On the contrary, the lower IBLSE group was found to be more sensitive to the effect of the ease-of-use of the system and paid more attention to the representational information quality (e.g. the consistency of the lectures). This outcome may suggest that learners with different IBLSE have diverse opinions about IBL activities.

While some studies have specialized in issues concerning TAM, others are concerned with the relationship between learners' perceived IBLSE and their satisfaction with IBL settings. As an instance, Artino (2008) surveyed 646 undergraduates and concluded that learners' IBLSE and their perception of the learning environment were significantly positive predictors of their satisfaction. Likewise, Johnson, Hornik, and Salas (2008) developed a model of e-learning effectiveness, which added social presence to other frequently studied variables, such as users' IBLSE, perceived usefulness, course interaction, and effectiveness. With an examination of survey responses from 345 individuals, they found that learners with higher IBLSE were more satisfied with the course than those with lower IBLSE.

Still others demonstrated the relationship between students' IBLSE and their learning outcomes in the IBL condition. For instance, to examine students' personal choices when taking Internet-based courses, Wang and Newlin (2002) investigated 122 college students and tested whether learners' IBLSE would predict their performance in the Internet-based sections of the class. The results showed that students' perceived IBLSE were predictive of their final exam scores. Moreover, students showing curiosity about the Internet-based program revealed higher IBLSE and had better class performance than those taking part in the course solely due to availability.

Similarly, Bolman *et al.* (2007) investigated the usability of a navigation support tool, which guided learners by generating advice on the next best step to take in a self-directed Internet-based course. Although they found that the navigation tool had not increased learners' IBLSE, it was indicated that learners with high IBLSE had completed more modules, adhered more often to the advice given, and were convinced that the navigation tool helped them plan the course. Therefore, it was suggested to incorporate IBLSE enhancing strategies in the navigational support of IBL activities.

Alteration of IBLSE

Among the reviewed papers, little research directly examined how IBLSE might be altered by certain types of IBL. Rather, these studies utilized "indirect" methods of investigation to reveal some potential avenues of fostering IBLSE. For instance, Bates and Khasawneh (2007) considered that evaluating the mediating role of IBLSE could provide a better understanding of the functional properties or potential enhancement of IBLSE and further clarify what factors might account for the differences among individuals in their participation in IBL activities. Accordingly, they proposed a mediated model to seek and to identify a number of theoretically based factors, which were believed to contribute to the development of IBLSE. On the basis of 288 university students' survey responses and self-reports, the research results revealed a partially mediated model, in which the block of antecedents (i.e. students' previous success with the IBL, instructor feedback, anxiety, pre-course training, and the perceived nature of IBL ability) had a direct effect on the dependent variables (i.e. students' outcome expectations, mastery perceptions, and the hours spent per week using the IBL technology to complete assignments for university courses) as well as an indirect effect through their influence on IBLSE. The finding was considered consistent with Bandura's (1982) premise that one of the strongest sources of self-efficacy beliefs is an individual's direct experience with the same or a similar phenomenon. It was proposed that instructional strategies, providing positive learning experiences with the IBL, may play a vital role in enhancing learners' IBLSE, fostering positive expectations, and encouraging their use of the technology.

With a sample of 223 learners taking part in an Internet-based program, Choi, Kim, and Kim (2007) confirmed that flow experience and attitude towards IBL had significant impacts on learners' IBLSE. Therefore, to enhance the effectiveness of IBL, it may not be sufficient to focus solely on learners' preferences; instead, to increase students' experience involvement, or intrinsic interest may be some possible ways to enhance learners' preferences, which may consequently contribute to students' IBLSE.

Besides, in an attempt to investigate students' conceptions of learning in the F2F condition, conceptions of IBL, and the differences between these conceptions, Tsai (2009) analyzed 83 Taiwanese college students' interview transcripts. The findings derived several categories of conceptions of F2F traditional-type learning and IBL, and it was suggested that the conceptions of IBL were often more sophisticated than those of F2F learning. In addition, learners' questionnaire responses revealed that the sophistication of the conceptions of IBL was associated with better searching strategies as well as higher IBLSE. These findings highlighted the need for fostering students' conceptions of learning by Internet-based environment, as they may enhance more sophisticated learning strategies and IBLSE.

Finally, different from previous research designs, one study was found to compare students' learning outcomes within a variety of classroom settings. Moneta and Kekkonen-Moneta (2007) assessed 414 students on not only affective learning (including intrinsic engagement, extrinsic engagement, and negative affect) but also IBLSE in an introductory computing course, which was taught once in a lecture format and twice in a rich interactive multimedia online format. IBLSE was assessed by a questionnaire item. The research results found that the IBL modules fostered more intrinsic engagement and higher IBLSE.

Summary of IBLSE research

In conclusion, while looking into the research relating to learners' IBLSE, probing their confidence in the participation and expected performance in the Internet-based activities, a great number of studies (e.g. Chang & Tung, 2008; Liaw, Huang, & Chen, 2007) were found to deal with the extended development of established models. For example, the notion of TAM is extensively applied in the relevant research. Moreover, some studies (e.g. Artino, 2008; Johnson, Hornik, & Salas, 2008) have discussed the relation between users' IBLSE and their satisfaction with IBL, whereas others (e.g. Bates & Khasawneh, 2007; Choi, Kim, & Kim, 2007; Tsai, 2009) were found to perceive IBLSE as a predictor or a mediator of students' learning outcomes in an Internet-based setting.

Indeed, by assessing students' IBLSE, researchers may have acquired indications about their expected outcomes derived from IBL activities. In fact, this finding may have also resulted from the fundamental feature of IBLSE, in which IBLSE was utilized to evaluate learners' perceptions of their learning in an Internet-based setting. Therefore, it may have been somehow unavoidable to associate IBLSE with the students' evaluation of their satisfaction or performance with regards to IBL.

Compared with the research found in the ISE and ASE&IBL categories, it was noted that relatively few studies (Moneta & Kekkonen-Moneta, 2007) had been conducted to compare learners' IBLSE among different learning environments. However, it was found that providing positive learning experiences with IBL activities (Bates & Khasawneh, 2007; Choi, Kim, & Kim, 2007) or the sophistication of the conceptions of IBL conceptions of IBL (Tsai, 2009) may have played an important role in enhancing learners' IBLSE.

Finally, several studies were found to probe learners' IBLES in specific domains in particular. For example, Chang and Tung (2008) and Tung and Chang (2008a, 2008b) examined students' IBLES in nursing contexts, whereas others investigated their IBLES in IBL of management (Lee & Lee, 2008), psychology (Wang & Newlin, 2002) or service academy (Artino, 2008). In general, the research results have indicated a positive influences of IBLES on either learners' intention, performance, or satisfactory toward IBL.

Discussions and Conclusions

State of self-efficacy research in IBL environments

Because of the increasingly important role self-efficacy plays in Internet-based learning (IBL), the relationship between self-efficacy and IBL has been widely investigated in the last decade. Consequently, the present study has collected and investigated 46 research papers published from 2000 to 2009 concerning these relations for a comprehensive literature review. Research is classified into three major categories: the Internet Self-Efficacy (ISE); the interplay between Academic Self-Efficacy and Internet-Based Learning (ASE&IBL); and the Internet-Based Learning Self-Efficacy (IBLSE).

Regarding the category of ISE research, the studies generally focus on the relationship between learners' Internet self-efficacy and learning processes. The relations among students' ISE and their attitudes, strategies, and preferences have been frequently examined. It is worth noting that because of the prolific development of Internet-based instruction, which has elicited various forms of IBL activities, research concerning ISE is likely to assess students' confidence in their skills or knowledge of operating specific Internet applications (such as communication) in IBL contexts instead of evaluating learners' ISE in general.

As for research regarding ASE&IBL, it is found that students' ASE is often applied to correlate with their performance, motivation, and perceptions of the effectiveness of Internet-based systems. In general, students' ASE has had positive effects on their academic outcomes resulted from IBL.

Finally, research on learners' IBLSE, which investigates learners' confidence in their expected performance in IBL, is discovered to mainly deal with the application of established models. For instance, a great amount of research has been arranged on the notion of Technology Acceptance Model (TAM). Moreover, IBLSE was found to be perceived as a predictor of students' learning outcomes and their satisfaction with IBL activities. In general, IBLSE was shown to have impacts on learners' satisfaction with IBL.

Evaluation of self-efficacy research in IBL environments

According to the original theory proposed by Bandura (1982, 1994), the source of self-efficacy is derived from multiple sources of efficacy information, including enactive mastery (e.g., past performance accomplishments resulting from previous experiences or training), verbal persuasion such as that resulting from collaboration and performance-related corrective feedback, and physiological arousal including changes in emotional states such as anxiety, fear, or positive anticipation. However, except for Francescato et al. (2006, 2007) and Johnson, Hornik, and Salas' (2008) reports concerning social persuasion, Bates and Khasawneh (2007), Yang, et al. (2008), and Chiou and Wan's (2007) investigation on mastery experiences, and Moneta and Kekkonen-Moneta's (2007) paper regarding affection arousal, relatively few empirical studies were found researching from the initially proposed concept of self-efficacy.

In addition, deriving from Moos and Azevedo's (2009) review of computer self-efficacy (CSE), three major findings were suggested: First, both learners' behavioral and psychological factors are related to CSE; secondly, CSE pertains to students' learning outcomes in computer-based learning (CBL) environments; and finally, CSE was found to be associated with users' navigational paths. In comparison with their findings, the current study found similar trends but somehow varied results in the research on self-efficacy in IBL environments. For instance, the results of research concerning the relationship between students' self-efficacy and their behavioral factors were found to be inconsistent in this current study. On one hand, some study findings revealed that ISE (Chu & Tsai, 2009; Lam & Lee, 2006; Yi & Hwang, 2003) or IBLSE (Chang & Tung, 2008; Liaw, Huang, & Chen, 2007) are related to students' subsequent use of IBL systems; on the other hand, others (Lu et al., 2007; Yang et al., 2007) stated that ISE had no significant effect on their following participation in IBL activities. This might imply there is a difference between the roles of CSE and ISE played in CBL and IBL environments respectively, indicating different natures could exist between CSE and ISE as well as between CBL and IBL.

Aside from the inconsistent findings about the relations between learners' self-efficacy and their behaviors, investigation outcomes on the relationship between students' self-efficacy and their psychological factors (such as perceived attitude, anxiety, and usefulness) seem to be consistent in this current study. Studies have, in general, indicated a positive relation between students' ISE and their attitude towards IBL (Peng, Tsai, & Wu, 2006; Torkzadeh, Chang, & Demirhan, 2006; Wu & Tsai, 2006) and a negative relation between learners' ISE and their perceived anxiety (Lam & Lee, 2006; Torkzadeh, Chang, & Demirhan, 2006; Yang et al., 2007).

Moreover, consistent with Moos and Azevedo's (2009) findings, the relationship between students' self-efficacy and their achieved outcomes in IBL was found to be positively correlated. For instance, Tsai and Tsai (2003) stated that students with higher ISE had better information search strategies and learnt better than their counterparts. Thompson, Meriac, and Cope (2002) found positive relations between students' self-efficacy (including both ISE and ASE) and the number of correct search results produced. In the ASE&IBL category, various researchers (Sins et al., 2008; Wang & Newlin, 2002; Yukselturk & Bulut, 2007) have claimed that ASE could serve as a positive predictor of

learners' final achievement in the IBL condition. Similarly, in the IBLSE category, Bolman et al. (2007) also revealed a positive relationship between learners' IBLSE and their IBL outcomes.

Furthermore, it was noted that when attempting to take measures probing learners' perceptions of self-efficacy, a significant amount of research (e.g. Chiou & Wan, 2007; Joo, Bong, & Choi, 2000; Thompson, Meriac, & Cope, 2002) has adopted *search tasks* to predict students' learning outcomes in the Internet-based setting because search tasks may still be considered as the most commonly implemented IBL activities. Nevertheless, in contrast to Moos and Azevedo's (2009) findings concerning CSE, no relevant research in this review was found on the relationship between individuals' perceived self-efficacy in the Internet-based setting and their navigational paths. More research may be needed to investigate on the relationship.

Finally, some methodological issues may be worthy of notice. First, it seems that all of the Internet-related research concerning self-efficacy is based on questionnaires or surveys for measuring self-efficacy. Researchers should find other ways of assessing students' Internet-related self-efficacy, such as interviews or observation. Most of the studies in this review employed a quantitative approach; qualitative or mixed research approaches are recommended for future research. In addition, among the 46 papers reviewed, 35 studies used university students as their samples. In other words, most of the participants invited to take part in the related studies were either undergraduates or graduate students in universities. It may be necessary to encourage learners with various kinds of demographic backgrounds to take part in the relevant research. Meanwhile, among the reviewed papers, 19 studies in America, 22 studies in Asia, and 5 studies in Europe were reported. Relevant research based on European samples is relatively rare. Finally, within all of the reviewed studies, only students' or employees' perceptions were probed; it may be interesting to investigate instructors' perceptions of Internet-related self-efficacy.

Future study

Because the papers selected for the current review were limited to those included in the SSCI database, other relevant research regarding self-efficacy may still be found to outline a more comprehensive review. Future studies can explore the differences between students' perceptions about CBL and IBL as well as compare students' CSE and ISE, simultaneously, in order to further examine the relationship between CSE and ISE. Secondly, researchers can further examine the construct of three categories of self-efficacy in IBL by assessing the significance and power of using ISE, ASE and IBLSE to predict students' IBL performances or outcomes. This could help researchers and educators realize more about the relationships among the three variables and learning outcomes in IBL environments. Thirdly, future studies can examine the relationship between students' perceived self-efficacy and their learning behaviors in specific Internet-based learning context such as the online search tasks mentioned above. Finally, more qualitative methods are suggested for future Internet-related self-efficacy assessments.

Acknowledgement

Funding for this research work is supported by the National Science Council, Taiwan, under grant numbers NSC 99-2511-S-011-005-MY3 and 100-2631-S-011-001.

References

(papers with * are those selected for current review)

- *Artino, A. R. (2008). Motivational beliefs and perceptions of instructional quality: Predicting satisfaction with online training. *Journal of Computer Assisted Learning*, 24, 260-270.
- Abdous, M., & Yoshimura, M. (2010). Learner outcomes and satisfaction: A comparison of live video-streamed instruction, satellite broadcast instruction, and face-to-face instruction. *Computers & Education*, 55, 733-741.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press.

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.
- *Bates, R., & Khasawneh, S. (2007). Self-efficacy and college students' perceptions and use of online learning systems. *Computers in Human Behavior*, 23, 175-191.
- Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P.A., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, 74, 379-439.
- *Bolman, C., Tattersall, W., Waterink, J., Janssen, B., van der Berg, R., van Es, R., & Koper, R. (2007). Learners' evaluation of a navigation support tool in distance education. *Journal of Computer Assisted Learning*, 23, 384-392.
- *Brown, S. W., Boyer, M. A., Mayall, H. J., Johnson, P. R., Meng, L., Butler, M. J., Weir, K., Florea, N., Hernandez, M. & Reis, S. (2003). The GlobalEd project: gender differences in a problem-based learning environment of international negotiations. *Instructional Science*, 34, 255-276.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, 100, 525-534.
- *Chang, S.C., & Tung, F. C. (2008). An empirical investigation of students' behavioural intentions to use the online learning course websites. *British Journal of Educational Technology*, 39, 71-83.
- *Chiou, W.-B. & Wan, C.-S. (2007). The dynamic change of self-efficacy in information searching on the Internet: Influence of valence of experience and prior self-efficacy. *The Journal of Psychology*, 141, 589-603.
- *Choi, D. H., Kim, J., & Kim, S. H. (2006). ERP training with a web-based electronic learning system: The flow theory perspective. *International Journal of Human-Computer Studies*, 65, 223-243.
- *Chu, R. J.-C., & Tsai, C.-C. (2009). Self-directed learning readiness, Internet self-efficacy and preferences towards constructivist Internet-based learning environments among higher-aged adults. *Journal of Computer Assisted Learning*, 25, 489-501.
- Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: development of a measure and initial test. *MIS Quarterly*, 19, 189-211.
- *Crippen K. J. & Earl, B. L., (2007). The impact of Web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. *Computers & Education*, 49, 809-821.
- Delialioglu, O., & Yildirim, Z. (2008). Design and development of a technology enhanced hybrid instruction based on MOLTA model: its effectiveness in comparison to traditional instruction. *Computers & Education*, 51, 474-483.
- Eastin, M. A., & LaRose, R. L. (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer Mediated Communication*, 6. Available at: <http://jcmc.indiana.edu/vol6/issue1/eastin.html> (October 23, 2010)
- *Farel, A., Umble, K., & Polhamus, B. (2001). Impact of an online analytic skills course. *Evaluation & the Health Professions*, 24, 446-459.
- *Francescato, D., Mebane, M.E., Porcelli, R., Attanasio C., & Pulino M. (2007). Developing professional skills and social capital through computer supported collaborative learning in university contexts. *International Journal of Human-Computer Studies*, 65, 140-152.
- *Francescato, D., Porcelli, R., Mebane, M., Cudetta, M., Klobas, J., & Renzi, P. (2006). Evaluation of the efficacy of collaborative learning in face to face and computer supported university contexts. *Computers in Human Behavior*, 22, 163-176.
- Girasoli, A. J., & Hannafin, R. D. (2008). Using asynchronous AV communication tools to increase academic self-efficacy. *Computers & Education*, 51, 1676-1682.
- Guan, Y.-H., Tsai, C.-C., & Hwang, F.-K. (2006). Content analysis of online discussion on a senior-high-school discussion forum of a virtual physics laboratory. *Instructional Science*, 34, 279-311.
- Hoffman, B., & Spatariu, A. (2008). The influence of self-efficacy and metacognitive prompting on math problem-solving efficiency. *Contemporary Educational Psychology*, 33, 875-893.
- *Hong, T. (2006). The Internet and tobacco cessation: The roles of Internet self-efficacy and search task on the information-seeking process. *Journal of Computer-Mediated Communication*, 11, 536-556.
- *Johnson, R. D., Hornik, S. & Salas, E. (2008). An empirical examination of factors contributing to the creation of successful e-learning environments. *International Journal of Human-Computer Studies*, 66, 356-369.
- *Joo, Y., Bong, M., & Choi, H.-J. (2000). Self-efficacy for self-regulated learning, academic self-efficacy, and Internet self-efficacy in Web-based instruction. *Educational Technology Research and Development*, 48, 5-17.
- Katz, J. E., & Aspeden, P. (1996, October). *Motivations for and barriers to Internet usage: Results of a national public opinion survey*. Paper presented at the 24th Annual Telecommunications Policy Research Conference, Solomons , MD .

- Karsten, R., & Roth, R. (1998). Computer self-efficacy: A practical indicator of student computer competency in introductory IS courses. *Informing Science*, 3, 61-68.
- *Kitsantas, A., & Chow, A. (2007). College students' perceived threat and preference for seeking help in traditional, distributed, and distance learning environments. *Computers & Education*, 48, 383-395.
- *Lam, J. C. I., & Lee, M. K. O. (2006). Digital inclusiveness-longitudinal study of Internet adoption by older adults. *Journal of Management Information Systems*, 22, 4, 177-206.
- *Lee, J. K. & Lee, W. K. (2008). The relationship of e-Learner's self-regulatory efficacy and perception of e-Learning environmental quality. *Computers in Human Behavior*, 24, 32-47.
- Lee, S. W.-Y., Tsai, C.-C., Wu, Y.-T., Tsai, M.-J., Liu, T.C., Hwang, F.-K., Lai, C.H., Liang, J.-C., Wu, H.-C., & Chang, C.-Y. (in press). Internet-based science learning: a review of journal publications. *International Journal of Science Education*.
- *Lee, Y. -C. (2006). An empirical investigation into factors influencing the adoption of an e-learning system. *Information & Management*, 30, 517-541.
- Lee, M. -C. (2010). Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation-confirmation model. *Computers & Education*, 54, 506-516.
- Li, N., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education*, 48, 301-317.
- *Liang, J.-C. & Tsai, C.-C. (2008). Internet self-efficacy and preferences toward constructivist Internet-based learning environments: A study of pre-school teachers in Taiwan. *Educational Technology and Society*, 11, 226-237.
- *Liaw, S. -S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, 51, 864-873.
- *Liaw, S.-S., Huang, H.-M., & Chen, G.-D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, 49, 1066-1080.
- *Lu, H. Y., Case, D. O., Lustria, M. L. A., Kwon, N., Andrews, J. E., Cavendish, S., & Floyd, B. R. (2007). Predictors of online information seeking by international students when disaster strikes their countries. *CyberPsychology & Behavior*, 10, 709-712.
- Marakas, G. M., Johnson, R. D., & Clay, P. F. (2007). The evolving nature of computer self-efficacy construct: An empirical investigation of measurement construction, validity, reliability and stability over time. *Journal of the Association for Information Systems*, 8, 16-46.
- Marakas, G. M., Yi, M. Y., & Johnson, R. D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. *Information Systems Research*, 9, 126-163.
- *Meyer, B. J. F., Middlemiss, W., Theodorou, E., Brezinski, K. L., McDougall, J., & Bartlett, B. J. (2002). Effects of structure strategy instruction delivered to fifth-grade children via the Internet with and without the aid of older adult tutors. *Journal of Educational Psychology*, 94, 486-519.
- *Moneta, G. B., & Kekkonen-Moneta, S. S. (2007). Affective learning in online multimedia and lecture versions of an introductory computing course. *Educational Psychology*, 27, 51-74.
- Moos, D. C., & Azevedo, R. (2009). Learning with computer-based learning environments: a literature review of computer self-efficacy. *Review of Educational Research*, 79, 576-600.
- Murphy, C.A., Coover, D., & Owen, S.V. (1989). Development and validation of the computer self-efficacy scale. *Educational and Psychological Measurement*, 49, 893-899.
- *Ong, C.-S. & Lai, J.-Y. (2006). Gender differences in perceptions and relationships among dominants of e-learning acceptance. *Computers in Human Behavior*, 22, 816-829.
- *O'Malley, M., & Kelleher, T. (2002). Papayas and pedagogy: Geographically dispersed teams and Internet self-efficacy. *Public Relations Review*, 28, 175-184.
- Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research*, 66, 543-578.
- Pajares, F. & Schunk, D. H. (2001). Self-beliefs and school success: self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *Perception* (pp. 239-266). London: Ablex Publishing.
- *Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society*, 12, 150-162.
- Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bullet in board discussions. *Computers & Education*, 42, 243-265.
- *Peng, H., Tsai, C.-C., & Wu, Y.-T. (2006). University students' self-efficacy and their attitudes toward the Internet: The role of students' perceptions of the Internet. *Educational Studies*, 32, 73-86.

- Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Ren, W-H (2000). Library instruction and college student self-sufficiency in electronic information searching. *Journal of Academic Librarianship*, 26, 323-328.
- Roca, J. C., Chiu, C. -M., & Martinez, F. J. (2006). Understanding e-learning continuance intention: An extension of the technology acceptance model. *International Journal of Human-Computer Studies*, 64, 683-696.
- *Schmidt, A. M. & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology*, 56, 405-429.
- Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 15-31). New York: Academic Press.
- *Sins, P. H. M., van Joolingen, W. R., Savelsbergh, E. R., & van Hout-Wolters, B. (2008). Motivation and performance within a collaborative computer-based modeling task: Relations between students' achievement goal orientation, self-efficacy, cognitive processing, and achievement. *Contemporary Education Psychology* 33, 58-77.
- *Tai, W. T. (2006). Effects of training framing, general self-efficacy and training motivation on trainees' training effectiveness, *Personnel Review*, 35, 51-65.
- Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of the research. *Review of Educational Research*, 76, 93-135.
- *Thompson, L. F., Meriac, J. P., & Cope, J. G. (2002). Motivating Online Performance: the influences of goal setting and Internet self efficacy. *Social Science Computer Review*, 20, 149-160.
- *Torkzadeh, G., Chang, J. -C., and Demirhan, D. (2006). A Contingency Model of Computer and Internet Self-efficacy. *Information and Management*, 43, 541-550.
- *Torkzadeh, G., & Van Dyke, T. P. (2001). Development and validation of an Internet self-efficacy scale. *Behaviour and Information Technology*, 20, 275-280.
- *Torkzadeh, G., & Van Dyke, T. P. (2002). Effects of training on Internet self-efficacy and computer user attitudes. *Computers in Human Behavior*, 18, 479-494.
- *Tsai, C.-C. (2009). Conceptions of learning versus conceptions of web-based learning: The differences revealed by college students. *Computers & Education*, 53, 1092-1103.
- *Tsai, M.-J., & Tsai, C.-C. (2003). Information searching strategies in web-based science learning: the role of Internet self-efficacy. *Innovations in Education and Teaching International*, 40, 43-50.
- *Tung, F. -C. & Chang, S. -C. (2008a). A new hybrid model for exploring the adoption of online nursing courses. *Nurse Education Today*, 28, 293-300.
- *Tung, F. -C. & Chang, S. -C. (2008b). Nursing students' behavioral intention to use online courses: A questionnaire survey. *International Journal of Nursing Studies*, 45, 1299-1309
- UNESCO (2003). Internet in education: Support material for educators. Available at: <<http://www.ite.ru/pics/publications/en/files/3214612.pdf>> (Oct 30, 2010)
- *Waldman, M. (2003). Freshmen's use of library electronic resources and self-efficacy. *Information Research*, 8, 1-27.
- *Wang, A. Y., & Newlin, M. H. (2002). Predictors of Web-Student Performance: the role of self- efficacy and reasons for taking an on-line class. *Computers in Human Behavior* 18, 151-163.
- White, B., & Frederiksen, J. (2005). A theoretical framework and approach for fostering metacognitive development. *Educational Psychologist*, 40, 211-223.
- *Wu, Y.-T., & Tsai, C.-C. (2006). University students' Internet attitudes and Internet self-efficacy: A study at three universities in Taiwan. *CyberPsychology & Behavior*, 9, 441-450.
- *Yang, H.-J., Lay, Y.-L., Tsao, W.-Y., Liou, Y.-C., & Lin, C.-K. (2007). The impact of language anxiety and self-efficacy on accessing Internet sites, *CyberPsychology and Behavior*, 10, 226-233.
- *Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model, *International Journal of Human-Computer Studies*, 59, 431-449.
- *Yukselturk, E. & Bulut, S. (2007). Predictors for student success in an online course. *Educational Technology & Society*, 10, 71-83.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59.

Comparison of Web 2.0 Technology Acceptance Level based on Cultural Differences

Sun Joo Yoo and Wen-hao David Huang

Department of Education Policy, Organization and Leadership, University of Illinois at Urbana-Champaign, IL, USA
// yoo23@illinois.edu // wdhuang@illinois.edu

ABSTRACT

In order to inform educators in higher education on the integration of Web 2.0 applications for engaging and effective learning experiences, this survey study compared the use and acceptance of Web 2.0 applications between American and Korean college students through the lens of cultural differences. Undergraduate students were recruited to participate in this study in the U. S. and the South of Korea, producing 183 usable responses. Targeting six Web 2.0 applications (blogs, instant messenger, online social communities/Facebook, online video sharing/YouTube, online video & audio conference/Skype, and social virtual communities/Second Life) the survey investigated five categories of technology acceptance based on the Unified Theory of Acceptance and Use of Technology. Significant differences on utilization level and the anxiety level for using them were found in numerous Web 2.0 applications. Korean students responded that most Web 2.0 applications are apprehensive for them to use when compared to their counterparts in the U.S.. The study further discussed the observed differences based on cultural theories and their implications on Web 2.0 learning technology integration.

Keywords

Web 2.0, cultural difference, technology preferences

Introduction

Today, college students use Web 2.0 applications more frequently than ever in and out of the classrooms. Applications such as blogs, instant messenger, online communities, video sharing tool, and web conferencing tool are gaining popularity. Students use them to create their own contents on the web, contribute and collaborate with others, and develop social networks via multiple formats of media and representation (O'Reilly, 2005). These activities although imply the possibility of using Web 2.0 applications for instructional purposes in higher education (Weller, Pegler, & Mason, 2005), it also raises the question of whether or not students can efficiently utilize the same Web 2.0 tools for learning (Huang & Behara, 2007). Even though students already use a variety of Web 2.0 applications on a daily basis, they may not know how to use them efficiently for gaining new knowledge or developing new skills. Educators in higher education interested in using Web 2.0 applications also need empirical evidences to help them integrate Web 2.0 applications in their instructional environments. This study aimed to explore these issues as an effort to promote the utilization of Web 2.0 applications to enhance learning experiences. Specifically, we were interested in exploring cultural differences' effect on Web 2.0 application utilization since research has suggested that such differences are present in the virtual world as well as in the physical world (Pfeil, Zaphiris, & Ang, 2006). Through comparison between participants from the U.S. and South Korea we aimed to investigate the usage of, and the technology acceptance level of Web 2.0 applications among college students. This study intended to answer the following questions:

- Do students from different countries prefer different Web 2.0 applications for learning?
- Are there different attitudes towards using Web 2.0 applications based on their cultural contexts?

Web 2.0 Applications for Learning

According to Anderson (2007), Web 2.0 is a collective term for a group of web-based technologies that broaden users' communication capabilities and options. Available tools include blogs, wikis, RSS feeds, online video sharing (e.g., YouTube, Google Video), and online social networking sites (eg., Facebook, LinkedIn, Ning). Timothy O'Reilly (2005), who initially conceived the term, defines Web 2.0 as an active and open web architecture that values users' proactive participation and contribution, which makes it more personalized and interactive than the previous generation Web applications (McLoughlin & Lee, 2007). Due to Web 2.0 applications' features in promoting proactive participation and collaborative sharing, they are suggested to bear great potential in educational settings.

In formal learning, Safran, Helic, & Gutzl (2007) emphasized that Web 2.0 applications make it possible to uphold critical and analytical thinking, facilitate intuitive and associational thinking, and support analogical thinking through ease access to rich information and interacting various opinions. Such applications might be efficient in conducting case studies due to their collaborative nature based on experiential learning approach (Huang & Behara, 2007). Furthermore, Web 2.0 applications enable users to connect to and collaborate with others with diverse interactions (Selwyn, 2007). For educators who implement student-centered learning activities and students who want to gain knowledge or skills through student-centered learning, Web 2.0 applications could also provide versatile opportunities. Boyd (2007) claimed that social aspects of Web 2.0 might support three activities that characterize student-centered learning. First is the support for conversational interaction. Second is the support for social feedback. Third is the support for social networks and relationship between people for enhancing the learning experience. *Facebook* (<http://www.facebook.com>), for example, is a social network community (SNC) that could enrich the learning process by allowing users to express themselves freely (Selwyn, 2007). Since *Facebook* has many useful facets for education such as reflective elements, peer-feedback with social context of learning (Selwyn, 2007), some educators have utilized its capacity for connecting students easily and optimistically (Lemeul, 2006). *Second Life*, an online virtual community, has also been used for educational purposes because it supports learning activities such as uploading personal opinions, participating in team work, and sharing knowledge and information made by users (Selwyn, 2007). Johnson & Levine (2008 & 2009) reported that in the next three to five years, higher education institutions will adopt a wide variety of Web 2.0 applications for the purposes of better engaging college students and enhancing instructional efficiencies. In terms of informal learning, Klammer, Chatti, Duval, Hummel, Hvannberg, Kravcik, Law, E., Naeve, & Scott, (2007) suggested that Web 2.0 applications could facilitate and enhance lifelong learning experience by connecting students in collaborative environments with diminishing boundaries around the world. People are engaged in a wide range of technologies-based informal learning at home and in the communities by continuously collaborating with others in interaction-rich social environments (Selwyn, 2007).

In sum, Web 2.0 applications, as discussed earlier, might be prominent to enable educators to create personalized, active, participatory, and cooperative learning environments (McLoughlin & Lee, 2007). In turn, educators can provide extensive opportunities for students who have various needs to enhance their learning experiences through enriched interactions and collaborations in Web 2.0 applications (Bryant, 2006; McLoughlin & Lee, 2007).

Web 2.0 Applications and Cultural Differences

The increasing use of the Internet in educational settings around the world has prompted discussions about the relationship between cultures and technologies (Holmes, 1998). Collis (1999) argued that culture is a critical factor in influencing how people accept, react to, and use the Internet. Recent studies have explored the cultural diversity of usage of the Internet. Li and Kirkup (2005) investigated cultural differences of the Internet use between students of universities in China and in the UK. Their results showed that British students used computers more for their study than Chinese students. Chinese students, however, had more confidence about their advanced computer skills. Allwood and Wang (1990) studied how students conceptualize computers in China and Sweden. They revealed that Swedish students were more pessimistic about the effects of computers on society than their Chinese counterparts. Other research (Brosnan & Lee, 1998; Collis & Williams, 1987) further concluded the effect of cultural differences on computer users' attitudes might be observable. Omar (1992) studied how college students have different attitudes towards computers and found that American students have more positive attitudes than Kuwaiti students.

According to Hofstede (1980, 1991), national differences can be understood in terms of national cultures. Cultures, in Hofstede's assertion, consist of four dimensions: power distance, individualism/collectivism, masculinity/femininity, and uncertainty avoidance. Power distance refers to "the degree of inequality in power between a less powerful individual and a more powerful other, in which individual and other belong to the same social system" (Hofstede, 2001, p. 83). High power distance of organizations tends to realize and accept that power is not given to employees equally (Hofstede, 1994, 1998). Individualism refers to the degree to which individuals feel they are "on their own" rather than part of a larger group identity. Masculinity refers to the degree to which a culture emphasizes competition, achievement, and "getting ahead" (Bearden, Money, & Nevins, 2006, p. 195). Uncertainty avoidance indicates the degree of tolerance for ambiguity and risk.

Among the four dimensions of cultural difference, power distance, individualism vs. collectivism, and uncertainty avoidance are particularly relevant in studying the use of Web 2.0 applications. Mandl (2009) explored comparing Chinese and German Blogs. They figured several results that Chinese expressed emotional and positive comments on Blogs compared to Germany. Users in Germany posted more negative comments on their blogs. This result assumes that users in China, whose culture is characterized by collectivism, tend to express less negative comments as they try not to say negative comments about others. A recent study explored the relations between the patterns of changes on wikis and the cultural background of the contributors, and suggested the tools themselves are neutral but cultural differences might affect how users utilize them (Pfeil, Zaphiris, & Ang, 2006). Their study found that respondents from high distance power countries tended to be reluctant to delete others' contents or links, even though they thought the content was incorrect. Respondents from high masculinity countries tended to add information and contribute to group products through lots of activities.

The most valuable advantage of Wiki is supporting collaboration among team members. According to Cogburn & Levinson (2003), one of the factors that hindered effective collaboration in virtual projects was the different communication styles from cultural differences. Gudykunst and Ting-Toomey (1988) reported the relationship between collectivism/individualism by Hofstede (1986) and high context/low context (Hall, 1976). In high context cultures, a lot of 'unspoken' meaning follows when users communicate with others. Hofstede (1984, 1986) addressed that North American and Western European countries tend to have individualism and low context culture. In contrast, South American and many Asian countries tend to be collectivistic and high context culture. Therefore, users in low culture cultures may not figure 'unspoken' meaning out in virtual communication. Guo, Tan, Turner, and Xu (2008) investigated preferences of communication media (face to face, telephone, email and instant messenger) from Australia and China. Chinese preferred to use telephone and instant messenger, while Australians preferred to use email. According to Hofstede (1989), China had an uncertainty index score that was lower than Australia. In order to avoid uncertainty, Australia may prefer to use email instead of telephone or instant messenger.

Flickr, a popular social photo-sharing tool, was investigated by Dotan and Zaphiris (2010). They revealed the cultural differences in websites driven by different countries' users such as Peru, Israel, Iran, Taiwan and the U. K. The users from Peru and Taiwan were less interested in sharing content compared to users from Iran and Israel since Peru and Taiwan had the highest power distance index and lowest individual index scores. However, the correlation between Hofstede's scores and quantitative data based on users' activities on websites with Flickr was weak. The gathered quantitative data may not reflect cultural differences or the index of cultural dimensions might be needed to collect.

Chapman and Lahav (2008) investigated social network sites from different cultures such as U.S, France, China, and South Korea. The results showed the differences in terms of the users' goals, typical pattern of self-expression, and common interaction behaviors. Users of popular social network sites in the U.S tend to publish more personal information. In contrast, users in China publish less personal information than those in the U.S.. In terms of self-expression, French users tend to discuss common topics instead of personal topics but Chinese users tend to discuss more personal topics. On interaction behaviors of social network sites, there are differences of user's behaviors between the U.S and South Korea. Shin (2010) investigated motivations for utilizing social networking sites from the U.S and South Korea. The results revealed that users from both countries believe social networking sites are useful and entertaining. Users in the U.S. utilize social networking sites because of extrinsic motivation but users in South Korea utilize them due to intrinsic motivation. In addition, users in South Korea connect with their real friends who they know and share content on Cy-world (Shin & Kim, 2008). However, users in the U.S. did not consider friends on MySpace as real friends (Dwyer, 2007). MySpace allows users in the U.S. to connect members who have the same interests or opinions and the social relations may be different between Cy-world and MySpace because of this reason (Shin, 2010). Fogg and Iizawa (2008) compared two popular social networking sites from the U.S and Japan to investigate how social networking tools might motivate users in different cultures. Facebook has direct and assertive motivators and Mixi, a popular social network site in Japan has indirect and subtle motivators.

Park, Mohan, and Ponnusamy (2009) investigated the learners' perception of Web 2.0 applications for learning. The results showed that Malaysian students learned more from a collectivist learning approach but the American students preferred an individualistic learning approach. However, learners in two countries responded that Web 2.0 applications were useful for learning. Eze (2009) conducted an online survey to explore which factors influence the utilization of Web 2.0 applications of European undergraduate and graduate students for learning. A total of 285 participants participated in the online survey from the Netherlands, Great Britain, and Ireland. This study reported

that culture is essential when users utilize Web 2.0 applications for learning. In particular, individualism, power distance, masculinity, and technology experience affected the utilization of Web 2.0 applications. Singh, Zhao and Hu (2003) compared Chinese and American websites focusing on cultural differences and stated, “the web is not a culturally neutral medium, but it is full of cultural markers that give country-specific websites a look and feel unique to the local culture”(p.63). That is, although the Internet is used around the globe, students use it differently based on their situational or national contexts (Li & Kirkup, 2007).

In sum, the usage of, and the technology acceptance level of Web 2.0 applications might be influenced by users' cultural differences since previous research has suggested that such differences are present in the virtual world. Therefore, this study focuses on examining cultural differences' effect on Web 2.0 application utilization through comparison between participants from the U.S. and South Korea.

Method

Participants

This survey study was conducted in 2008 at a public midwestern university in the U.S. and at two private universities in South Korea. A total of 314 students (*107 American students and 207 Korean students*) were recruited to participate, producing 183 usable responses (*83 American and 100 Korean*). Comparison analyses were conducted to identify the difference between the two sets of survey data. All data were collected via voluntary participation by a web-based survey interface.

Measuring Utilization Level of Web 2.0 Applications

To cover a wide range of technologies that have been utilized in e-learning settings in higher education, this study targeted six Web 2.0 applications: blogs, instant messenger, online social communities (eg., Facebook), online video sharing (eg., YouTube), online video & audio conferencing tools, and social virtual environment (eg., Second Life). To investigate the utilization level of Web 2.0 applications, the research team created the survey based on selected categories of *Unified Theory of Acceptance and Use of Technology* (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003), which includes performance expectancy, effort expectancy, attitude, social influence, and anxiety. See Table 1 for survey items on Web 2.0 utilization. UTAUT is the synthesis of eight other models (ie., theory of reasoned action, technology acceptance model, motivational model, theory of planned behavior, model of PC utilization, innovation diffusion model, and the social cognitive theory). The instrument measures a unified technology acceptance rate expressed by individuals or organizations.

Table 1. Acceptance of Web 2.0 applications

| Categorization | Questions | Code |
|--|--|------|
| Performance expectancy | 1. I would find it useful in my learning tasks | 1 |
| | 2. Using it enables me to accomplish tasks more quickly | 2 |
| | 3. Using it increases my productivity | 3 |
| Effort expectancy | 4. Learning to use it is easy for me | 4 |
| Attitude toward using Web 2.0 applications | 5. Using it for learning is a good idea | 5 |
| | 6. It makes learning more interesting | 6 |
| | 7. I like learning with it | 7 |
| Social influence | 8. People who influence my behavior think that I should use it | 8 |
| Anxiety | 9. I feel apprehensive about using it | 9 |
| | 10. It is somewhat intimidating to me | 10 |

Procedures of the Study

The survey was developed using an online survey tool, Survey Monkey (<http://www.surveymonkey.com>), which can be accessed from anywhere via the Internet. Participants had as much time as they needed to complete the

survey. But they can access the online survey once. Questions were based on a 7-point Likert scale with response options ranging from *strongly disagree* (1) to *strongly agree* (7). Demographic information regarding students was also obtained. Participants were recruited from undergraduate and graduate programs via random selections. The researchers then contacted course instructors for permissions to post the research information in their classrooms explaining the purpose and procedure of the study. The online survey was available for four months before the data analysis.

Results

Based on the research questions, this study used both descriptive and inferential statistics for data analysis. The research team first used descriptive statistics to report the distribution of Web 2.0 utilization levels. Inferential statistics (ie., *t*-test) was then conducted to identify the difference on participants' Web 2.0 utilization levels between samples from the two countries.

In total, 107 students from a public mid-western university in the U.S. and 207 students from two private universities in South Korea participated in the study. However, 131 participants' datasets were removed due to incompleteness or errors. The research team analyzed 183 usable responses for the utilization of Web 2.0 applications (59% of 314 submitted responses). Of the 183 completed surveys, 75 of them were male (40.9%) and 77 were female (42.0%). The data consists of 18 (9.8%) freshmen, 45 (24.5%) sophomores, 50 (27.3%) juniors, 51 (27.8%) seniors, and 14 (7.6%) graduates. The respondents also reported their academic majors. See Table 2 for the demographic and academic major data.

Table 2. Descriptive statistics of the respondents (%)

| Demographics | U.S (n=83) | South Korea (n=100) | Total (n=183) |
|------------------------------|-----------------------|--------------------------------|--------------------------|
| Sex | | | |
| Male | 42(50.6%) | 33(33%) | 75(40.9%) |
| Female | 21(25.3%) | 56(56%) | 77(42.0%) |
| No response | 20(24.0%) | 11(11%) | 31(16.9%) |
| Year | | | |
| Freshman | 17(20.4%) | 1(1%) | 18(9.8%) |
| Sophomore | 9(10.8%) | 36(36%) | 45(24.5%) |
| Junior | 10(12.0%) | 40(40%) | 50(27.3%) |
| Senior | 32(38.5%) | 19(19%) | 51(27.8%) |
| Graduate | 10(12.0%) | 4(4%) | 14(7.6%) |
| No response | 5(6%) | N/A | 5(2.7%) |
| Major | | | |
| Business/Management | 10(12.0%) | 7(7%) | 17(9.2%) |
| Engineering/Computer Science | 22(26.5%) | 7(7%) | 29(15.8%) |
| Education | 15(18.0%) | 61(61%) | 76(41.5%) |
| Liberal Arts/Social Science | 16(19.2%) | 19(19%) | 35(19.1%) |
| Other | 12(14.4%) | 6(6%) | 18(9.8%) |
| No response | 8(9.6%) | N/A | 8(4.3%) |

Descriptive statistical analysis

The usage frequency of Web 2.0 applications from 183 participants was identified (see Table 3). There were 37% American students responded that they did not know what a blog is. On the contrary, 34% Korean students responded that they used blogs more than 7 times per week. Online video & audio conferencing tool and social virtual environments are mostly unfamiliar to participants. The data indicated that 52% of American students and 48% of Korean students did not know what an online video & audio conferencing tool is. In addition, 78% of American and 66% of Korean respondents reported that they did not know what a social virtual environment is.

Table 3. Analysis of usage frequency of Web 2.0 applications per week

| Web 2.0 applications | Country | I don't know what this is | 1 time | 2 - 4 times | 5 - 6 times | More than 7 times |
|--|---------|---------------------------|---------|-------------|-------------|-------------------|
| Blogs | U.S. | 31(37%) | 31(37%) | 9(11%) | 2(2%) | 10(12%) |
| | S.K. | 1(1%) | 26(26%) | 17(17%) | 22(22%) | 34(34%) |
| Instant Messenger | U.S. | 2(2%) | 20(24%) | 14(17%) | 7(8%) | 40(48%) |
| | S.K. | 2(2%) | 14(14%) | 21(21%) | 22(22%) | 41(41%) |
| Online social communities (e.g., Facebook) | U.S. | 10(12%) | 9(11%) | 13(16%) | 12(14%) | 39(47%) |
| | S.K. | 5(5%) | 25(25%) | 36(36%) | 18(18%) | 15(15%) |
| Online video sharing (e.g., YouTube) | U.S. | 7(8%) | 24(29%) | 26(31%) | 10(12%) | 16(19%) |
| | S.K. | 4(4%) | 47(47%) | 27(27%) | 10(10%) | 12(12%) |
| Online video & audio conferencing tool (e.g., Skype) | U.S. | 43(52%) | 29(35%) | 5(6%) | 3(4%) | 3(4%) |
| | S.K. | 48(48%) | 39(39%) | 10(10%) | 1(1%) | 2(2%) |
| Social virtual environment (e.g., Second Life) | U.S. | 65(78%) | 12(14%) | 2(2%) | 2(2%) | 0(0%) |
| | S.K. | 66(66%) | 18(18%) | 11(11%) | 2(2%) | 3(3%) |

Table 4 shows the technology acceptance of Web 2.0 for learning on all six Web 2.0 applications. In average, all participants had a more positive attitude towards using instant messenger and online video sharing than other Web 2.0 tools. While American participants found that instant messenger and online video sharing are the easiest to use, Korean participants had very positive attitudes toward using blogs for their learning. In average, participants felt intimidated using a social virtual environment for learning. The anxiety level associated with using all Web 2.0 tools, however, is relatively low on a 7-point Likert scale.

Table 4. Analysis of the technology acceptance of Web 2.0

| | Country | Blogs | Instant Messenger | Online social communities | Online video sharing | Online video & audio conferencing tool | Social virtual environments |
|--------------------------|--|-------|-------------------|---------------------------|----------------------|--|-----------------------------|
| Performance | I would find it useful in my learning tasks | U.S. | 3.49 | 4.23 | 3.42 | 4.93 | 3.61 |
| | | S.K. | 4.50 | 4.15 | 4.46 | 4.30 | 3.82 |
| | Using it enables me to accomplish tasks more quickly | U.S. | 3.20 | 4.22 | 3.23 | 4.19 | 3.65 |
| | | | 4.60 | 4.27 | 4.43 | 4.29 | 3.74 |
| | Using it increases my productivity | U.S. | 3.10 | 3.75 | 2.80 | 4.00 | 3.53 |
| | | | 4.40 | 4.00 | 4.43 | 4.21 | 3.79 |
| Effort Expectancy | Learning to use it is easy for me | U.S. | 4.33 | 5.69 | 4.92 | 5.22 | 3.89 |
| | | S.K. | 4.41 | 4.42 | 4.57 | 4.40 | 3.70 |
| Attitude | Using it for learning is a good idea | U.S. | 3.86 | 4.45 | 3.51 | 4.40 | 3.65 |
| | | | 4.74 | 4.47 | 4.68 | 4.37 | 4.03 |
| | It makes learning more interesting | U.S. | 3.81 | 4.28 | 3.70 | 4.90 | 3.52 |
| | | | 4.44 | 4.13 | 4.47 | 4.04 | 3.86 |
| | I like learning with it | U.S. | 3.57 | 4.16 | 3.67 | 4.72 | 3.43 |
| | | | 4.19 | 3.87 | 4.33 | 3.96 | 3.55 |

| | | | | | | | | |
|--|--|------|------|------|------|------|------|------|
| Social Influence | People who influence my behavior think that I should use it. | U.S. | 2.95 | 3.70 | 3.87 | 4.20 | 3.39 | 2.96 |
| | | S.K. | | 3.43 | 3.61 | 4.03 | 3.81 | 3.38 |
| Anxiety | I feel apprehensive about using it | U.S. | 2.62 | 2.60 | 2.73 | 2.52 | 2.86 | 2.92 |
| | | S.K. | | 4.78 | 5.14 | 4.77 | 4.66 | 3.90 |
| It is somewhat intimidating to me | U.S. | 2.31 | 1.98 | 2.22 | 2.31 | 2.90 | 2.77 | |
| | S.K. | | 2.62 | 2.43 | 2.71 | 2.83 | 3.18 | 3.18 |

Table 5 shows the comparison of technology acceptance of Web 2.0 applications for learning from students in the U.S. and South Korea based on a *t*-test. Korean students had a more positive attitude towards using blogs compared to American students in terms of performance, attitude, social influence, and anxiety perspectives. In particular, Korean students responded that using blogs is a good idea as it makes learning more interesting. American students responded that learning to use instant messenger is easy, and it is not intimidating when compared to Korean students. In terms of performance and attitude categories of the technology acceptance, Korean students' responses in using online communities for learning are more positive than their American counterparts.

American students had a more positive attitude towards using online video sharing than Korean students. They felt that using online video sharing makes learning more interesting and they like learning with online video sharing. They felt that using online video sharing is more useful in their learning tasks. American participants responded that using online video sharing is very easy for them. They also felt less intimidated when compared to Korean respondents. The analysis further indicated a significant difference between American and Korean students in the anxiety category for all six Web 2.0 applications. Korean participants felt more anxious about using those applications than American participants. Korean respondents felt apprehensive about using online audio & video conference tool. Korean students who had experience using social virtual communities had positive attitudes compared to American respondents. In terms of performance, Korean participants responded that social virtual communities are very useful and increase their productivity. They also felt that using social virtual communities enables them to accomplish tasks more quickly when compared to American participants.

Table 5. Comparison of the technology acceptance (*t*-test)

| Web 2.0 applications | | <i>t</i> -test for Equality of Means | | | | |
|--------------------------|------------------|--------------------------------------|-------|--------------------|-----------------|-----------------------|
| Technology acceptance | Question | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Blogs | Performance | 1 | -3.76 | 133.29 | 0.00** | -1.01 |
| | | 2 | -5.47 | 146.83 | 0.00** | -1.40 |
| | | 3 | -5.38 | 141.01 | 0.00** | -1.30 |
| | Effort | 4 | | | | |
| | Expectancy | | -0.28 | 118.35 | 0.78 | -0.08 |
| | Attitude | 5 | -3.34 | 124.59 | 0.00** | -0.88 |
| | | 6 | -2.49 | 116.52 | 0.01** | -0.63 |
| | | 7 | -2.30 | 121.08 | 0.02* | -0.62 |
| | Social influence | 8 | | | | |
| | | | -1.93 | 131.88 | 0.06 | -0.48 |
| | | 9 | -9.41 | 144.33 | 0.00** | -2.17 |
| | Anxiety | 10 | -1.32 | 148.84 | 0.19 | -0.31 |
| Instant messenger | Performance | 1 | 0.34 | 142.41 | 0.74 | 0.08 |
| | | 2 | -0.22 | 142.64 | 0.83 | -0.05 |
| | | 3 | -1.04 | 130.39 | 0.30 | -0.25 |
| | Effort | 4 | | | | |
| | Expectancy | | 5.09 | 140.15 | 0.00** | 1.27 |
| | Attitude | 5 | -0.11 | 141.56 | 0.91 | -0.02 |
| | | 6 | 0.63 | 127.14 | 0.53 | 0.15 |

| | | | | | | | |
|--|------------------|----|--------|--------|--------|-------|------|
| | | 7 | 1.16 | 137.73 | 0.25 | 0.29 | 0.25 |
| | Social influence | 8 | | 0.35 | 141.53 | 0.73 | 0.09 |
| | Anxiety | 9 | -11.38 | 150.06 | 0.00** | -2.54 | 0.22 |
| | | 10 | -2.31 | 181.00 | 0.02* | -0.45 | 0.20 |
| Online social communities | Performance | 1 | -4.62 | 124.98 | 0.00** | -1.04 | 0.22 |
| | | 2 | -5.23 | 133.58 | 0.00** | -1.20 | 0.23 |
| | | 3 | -7.40 | 139.13 | 0.00** | -1.63 | 0.22 |
| | Effort | 4 | | | | | |
| | Expectancy | | 1.26 | 112.78 | 0.21 | 0.35 | 0.27 |
| | Attitude | 5 | -5.09 | 122.08 | 0.00** | -1.17 | 0.23 |
| | | 6 | -3.53 | 136.79 | 0.00** | -0.77 | 0.22 |
| | | 7 | -2.72 | 131.20 | 0.01** | -0.66 | 0.24 |
| | Social influence | 8 | - | 135.45 | 0.53 | -0.16 | 0.26 |
| | Anxiety | 9 | -8.85 | 141.11 | 0.00** | -2.04 | 0.23 |
| | | 10 | -2.23 | 181.00 | 0.03* | -0.49 | 0.22 |
| Online video sharing | Performance | 1 | 2.76 | 155.73 | 0.01** | 0.63 | 0.23 |
| | | 2 | -0.41 | 144.36 | 0.68 | -0.10 | 0.24 |
| | | 3 | -0.90 | 139.90 | 0.37 | -0.21 | 0.23 |
| | Effort | 4 | | | | | |
| | Expectancy | | 3.27 | 136.27 | 0.00** | 0.82 | 0.25 |
| | Attitude | 5 | 0.11 | 142.05 | 0.91 | 0.03 | 0.25 |
| | | 6 | 3.62 | 136.50 | 0.00** | 0.86 | 0.24 |
| | | 7 | 2.99 | 132.83 | 0.00** | 0.76 | 0.26 |
| | Social influence | 8 | 1.77 | 181.00 | 0.08 | 0.39 | 0.22 |
| | Anxiety | 9 | -9.41 | 159.29 | 0.00** | -2.14 | 0.23 |
| | | 10 | -2.23 | 181.00 | 0.03* | -0.52 | 0.23 |
| Audio & Video conferencing tool | Performance | 1 | -0.82 | 139.31 | 0.42 | -0.21 | 0.25 |
| | | 2 | -0.37 | 181.00 | 0.71 | -0.09 | 0.24 |
| | | 3 | -1.06 | 146.10 | 0.29 | -0.26 | 0.25 |
| | Effort | 4 | | | | | |
| | Expectancy | | 0.71 | 138.23 | 0.48 | 0.19 | 0.27 |
| | Attitude | 5 | -1.50 | 147.13 | 0.14 | -0.38 | 0.25 |
| | | 6 | -1.38 | 141.42 | 0.17 | -0.34 | 0.25 |
| | | 7 | -0.47 | 135.90 | 0.64 | -0.12 | 0.25 |
| | Social influence | 8 | 0.02 | 150.29 | 0.98 | 0.01 | 0.25 |
| | Anxiety | 9 | -4.37 | 166.83 | 0.00** | -1.04 | 0.24 |
| | | 10 | -1.11 | 167.04 | 0.27 | -0.28 | 0.25 |
| Social virtual Communities | Performance | 1 | -2.40 | 148.94 | 0.02* | -0.60 | 0.25 |
| | | 2 | -2.25 | 146.17 | 0.03* | -0.54 | 0.24 |
| | | 3 | -2.34 | 147.68 | 0.02* | -0.57 | 0.25 |
| | Effort | 4 | | | | | |
| | Expectancy | | -1.06 | 138.60 | 0.29 | -0.28 | 0.27 |
| | Attitude | 5 | -2.71 | 146.92 | 0.01** | -0.69 | 0.25 |
| | | 6 | -2.17 | 141.20 | 0.03* | -0.55 | 0.25 |
| | | 7 | -1.44 | 143.17 | 0.15 | -0.36 | 0.25 |
| | Social influence | 8 | -0.87 | 148.76 | 0.38 | -0.22 | 0.25 |
| | Anxiety | 9 | -2.64 | 150.06 | 0.01** | -0.64 | 0.24 |
| | | 10 | -1.60 | 181.00 | 0.11 | -0.41 | 0.26 |

Note: *p<0.05 **p<0.01

Discussion

This study has shown that American and South Korean students tend to differ in their technology acceptance levels and the usage of Web 2.0 applications for learning. Korean students reported positive attitudes towards using blogs and participating in online social communities but they had high anxiety levels in using online conferencing tools (eg., Skype) and social virtual environments (eg., Second Life). American students perceived a high difficulty level in using several Web 2.0 applications, such as social virtual environment tools, while they felt at ease in participating in online social communities (eg., Facebook). American students felt optimistic in using instant messenger and online video sharing for learning but their survey responses showed lower levels of anxiety towards online conferencing and social virtual environments.

According to Communications Workers of America (CWA, 2009), South Korea has the fastest average Internet connection speed and the highest rate of broadband connectivity (Akamai Technologies, 2008) in the world. This technical fact, however, does not seem to support both descriptive and inference analyses of survey results since Korean students felt more anxious when using Web 2.0 applications than American students. Moreover, they were intimidated by instant messenger and online conferencing tools (eg., Skype). Regardless of which Web 2.0 applications, the study showed that Korean students are more apprehensive in using them for learning than American students.

The obvious place to look for an explanation of the differences in Web 2.0 acceptance and utilization resides in two different cultural contexts of U.S. and South Korea. Sociologists have provided useful cultural dimensions focusing on attitudes towards power distance, individualism/collectivism, and uncertainty avoidance tendencies. According to Hofstede (1994, 1998), high power distance implies that the lower power members of organizations tend to realize and accept that power is not given to them equally. Korean students responded that participating in online social communities or social virtual communities is helpful toward their learning, but the results showed that they did not use them often in comparison to American students. Due to the perceived power distance, Korean students might hesitate to share their own thoughts through online social communities even though they thought that sharing different perspectives or opinions is useful for learning. Blogs, which were preferred by Korean students, have different characteristics in comparison to online social communities. Blogs can only be updated by the author, which insulates the blog author from the effect of power distance and allows them to express his thoughts or opinions freely. Blogs might be the better place for Korean students to articulate their opinions than other Web 2.0 applications because they do not need to worry about potential criticism by other participants who might be placed at a higher level of the power hierarchy.

The second cultural difference can be inferred from communication styles (Hall, 1979). According to Hofstede (1991), South Korean is very high on their collectivism scores while Americans in general have very high individualism scores. Gudykunst and Ting-Toomey (1998) contended that individualistic cultures were associated with low-context communication while the collectivistic cultures were associated with high-context cultures. Koreans, therefore, might prefer face-to-face contacts to their counterparts in Western cultures (Triandis, 1994). People who are from individualistic cultures tend to focus on themselves as unique entities, but people from collectivistic cultures see themselves as members of a group (Triandis, 1988). In high-context cultures, meanings in communication are derived from the situation, involved social relationships, and the content. People in individualistic cultures, on the other hand, rely far less on contexts to interpret meanings of communications. They tend to focus on the content of the communication. This might explain why American students felt that using instant messenger is very easy, and do not feel intimidated by using it as IM satisfies the need to communicate on the content only. On the contrary, Korean students responded that they were intimidated by using instant messengers. They may feel that they should understand the context as well as the content of the communication. Communicating through instant messengers without a clear context might not be formal enough for Koreans, which might also lead to miscommunication due to the lack of contextual cues.

The third cultural difference may be caused by uncertainty avoidance. Uncertainty avoidance in a society is often reflected in formal educational systems. Students who are in strong uncertainty avoidance cultures prefer structured learning situations with well-organized objectives, timetables, and assignments (Hofstede, 2001). When uncertainty avoidance is weak such as in Britain or U.S., students and instructors both care less about the structure of the learning process. These low uncertainty avoidance cultures, for example, encourage students to participate in online

discussions freely. So Korean students who belong to the culture of high uncertainty avoidance might hesitate to participate in online social communities or online virtual communities.

The impact of different cultures may bring conflicts for instructors and learners in traditional learning settings that are isolated from the rest of the world (Uzuner, 2009). However, it is common to encounter students with different cultural backgrounds in learning environments that are connected by the Internet. Thus it is very important to understand the significant differences between two different cultures toward the acceptance and utilization of Web2.0 applications for learning. This study could provide practical insights for educators on how to effectively integrate Web 2.0 applications for learners coming from various cultural backgrounds (Uzuner, 2009).

The findings of this study imply that instructors in culturally diverse instructional settings should consider the relationship between cultural differences and the usage and acceptance of Web 2.0 applications to facilitate students' learning activities. Instructors also need to consider both learning objectives of the communication and cultural attitudes toward Web 2.0 applications. One way might be to intentionally choose specific Web 2.0 application for students who are reluctant to use them, as they tend to stick to specific tools based on their cultural preferences. For example, generally instant messenger is an effective tool to interact flexibly among team members. However, if the objective is to improve formal expressions of communication, instant messenger might not be a good tool for it. If the purpose is to facilitate active interactions among students, instructors should select instant messenger even though students who are from high power distance countries are not familiar with using instant messengers. Using wikis for collaborative interactions among students is another example. Students who are from high power distance countries may not comfortable contributing to wikis because they think they may have no right to delete or correct the content contributed by others. To support students' collaboration, instructors can choose wikis even though they realize students' cultural differences. In the case of unwilling students, instructors could help students to understand the purpose behind the use of the chosen Web 2.0 applications and challenge them to overcome their cultural attitudes towards the technology. The other approach might be choosing several Web 2.0 applications based on students' cultural differences. Students who have different cultural backgrounds might select different tools for their learning based on their cultural preferences.

To summarize, cultural difference may influence how students accept technology and how to use it for learning. With the inevitable trend of organizations utilizing Web 2.0 applications for educational and professional purposes, cultural issues should become an important topic. Although Web 2.0 applications themselves are neutral but users are always affected by their cultural contexts. Therefore, understanding cultural differences and their potential impact could help educators to efficiently integrate appropriate Web 2.0 tools in educational settings.

Limitations and Future Research

Given the small sample size, we are unable to generalize our findings to different instructional settings. The preliminary results, however, offer two basic principles on which to base further research. First, students from different cultural contexts do perceive and utilize Web 2.0 applications differently for learning purposes. Second, it is not difficult to see classes that include and mix students from variety of cultural backgrounds, so it is valuable to examine how such students accept, react to, and use Web 2.0 applications for collaboration among team members. Third, potential differences of Web 2.0 utilization in instructional settings are very valuable for educators. To further investigate these principles, the research team intends to conduct in-depth task analyses for all six Web 2.0 applications in similar cultural comparison settings, and align them with intended learning outcomes. As a result we will be able to optimize learners' performance in learning environments enriched by Web 2.0 applications.

References

- Anderson, P. (2007). What is Web 2.0? Ideas, technologies and implications for education. *Technical report*, JISC.
- Allwood, C., and Wang, Z-M.(1990). Conceptions of computers among students in China and Sweden. *Computers in Human Behavior*, 6, 185-199.
- Akamai Technologies (2008). 4th Quarter 2008. *The State of the Internet*, 1(4).

- Bearden, W. O., Money, R. B., Nevins, & J. L. (2006). Multidimensional versus unidimensional measure in assessing national culture values: The Hofstede VSM 94 example. *Journal of Business Research*, 59, 195-203.
- Boyd, D. (2007). The significance of social software. In T. N. Burg & J. Schmidt (Eds.), *blogs Talks reloaded: Social software research & cases* (pp. 15-30). Norderstedt, Germany: Books on Demand.
- Brosnan, M., & Lee, W. (1998). A cross-cultural comparison of gender differences in computer attitudes and anxieties: the United Kingdom and Hong Kong. *Computers in Human Behavior*, 14(4), 559-577.
- Bryant, T. (2006). Social software in Academia. *Educause Quarterly*, 2, 61-64.
- Chapman, C. N., & Lahav, M. (2008). International Ethnographic Observation of Social Networking Sites. *Paper presented at the CHI 2008*, April 5–10, Florence, Italy.
- Chau, P. Y. K. (2008). Cultural differences in diffusion, adoption, and infusion of Web 2.0. *Journal of Global Information Management*, 16(1), I-III. 6.
- Cogburn, D. L., & Levinson, N. S. (2003). US-Africa virtual collaboration in globalization studies: Success factors for complex, cross-national learning teams. *International Studies Perspectives*, 4, 34-51.
- Collis, B. (1999). Designing for differences: cultural issues in the design of WWW-based course-support sites. *British Journal of Educational Technology*, 30(3), 201-215.
- Collis, B. A., & Williams, R. L. (1987). Cross-cultural comparison of gender differences in adolescent attitudes toward computers and selected school subjects. *Journal of Educational Research*, 81(1), 17-27.
- Communications Workers of America (2009). *Nationwide Study of Real-Time Internet Connection Speeds Shows U.S. Still Lags behind Other Advanced Nations, Posts Small Gains Last Two Years*. Retrieved November 22, 2009, from <http://www.cwa-union.org/news/nationwide-study-of-real-time-internet-connection-speeds-shows-u-s-still-lags-behind-other-advanced-nations-posts-small-gains-last-two-years.html>.
- Dotan, A., & Zaphiris, P. (2010). A Cross-Cultural Analysis of Flickr Users from Peru, Israel, Iran, Taiwan and the United Kingdom. *International Journal of Web Based Communities*, 6(3), 284-302.
- Dwyer, C. (2007). Digital relationships in the MySpace generation: results from a qualitative study. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, New York: IEEE, 19.
- Eze, E.O. (2009) *Making sense of Web 2.0 technology: do European students use the social media applications for educational goals?* Twente: University of Twente.
- Gudykunst, W. B., & Ting-Toomey, S. (with Chua, E.) (1998). *Culture and Interpersonal Communication*. Newbury Park: Doubleday.
- Guo, Z., Tan, F. B., Turner, T., & Huizhong Xu. (2008). An exploratory investigation into instant messaging preferences in two distinct cultures. *IEEE Transactions on Professional Communication*, 51(4), 396-415.
- Hall, E. T. (1976). *Beyond culture*. Garden City, NY: Doubleday & Company.
- Harris, M. & Johnson, O. (2003). *Cultural anthropology* (6th Ed.). Allyn and Bacon.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
- Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of Management*, 1(2), 81-99.
- Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations*, 10, 301-320.
- Hofstede, G. (1989). Organising for cultural diversity. *European Management Journal*, 7(4), 390-397.
- Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.
- Hofstede, G. (1994). *Value Survey Module 1994 Manual*. Maastricht: Institute for Research on Intercultural Cooperation.
- Hofstede, G. (1998). A case for comparing apples with oranges: International differences in values. *International Journal of Comparative Sociology*, 39(1), 16-31.
- Hofstede, G. (2001). *Culture's consequences*. Thousand Oaks, CA: Sage.
- Holmes, (1998). Cross-cultural differences of use of information technology in education: a comparative study of the use of computers in Japanese and British classrooms. *Unpublished PhD dissertation*, Wolfson College, University of Cambridge.
- Huang, C.D., & Behara, R.S. (2007). Outcome-driven experiential learning with Web 2.0. *Journal of Information Systems Education*, 18(3), 329-36.
- Johnson, L., Levine, A., & Smith, R. (2008). *The 2009 Horizon Report*. Austin, Texas: The New Media Consortium.

- Johnson, L., Levine, A., & Smith, R. (2009). *The 2009 Horizon Report*. Austin, Texas: The New Media Consortium.
- Klamma, R., Chatti, M.A., Duval, E. Hummel, H., Hvannberg, E. H., Kravcik, M., Law, E., Naeve, A., & Scott, P. (2007). Social software for life-long learning. *Journal of Educational Technology and Society*, 10(3), 72-83.
- Lemeul, J. (2006). Why I registered on Facebook. *The Chronicle of Higher Education*, 53 (2), retrieved May 1, 2011 from <http://chronicle.com/weekly/v53/i02/02c00101.htm>.
- Li, N., & Kirkup, G. (2007). Gender and cultural differences in internet use: A study of China and the UK. *Computers & Education*, 48(2), 301-317.
- Mandl, T. (2009). Comparing Chinese and German Blogs. *Paper presented at the 20th ACM Conference on Hypermedia and hypermedia*, June 29 - July 1, Torino, Italy.
- McLoughlin, C. & Lee, M.J.W. (2007). Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era. *Proceedings of ASCILITE Conference 2007*(pp.664- 673), Singapore.
- O'Reilly, T. (2005). *What is Web 2.0: design patterns and business models for the next generation of software*. Retrieved on April 26, 2011 from <http://www.oreilly.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
- Omar, M. H. (1992). Attitudes of college students towards computers: a comparative study in the United States and the Middle East. *Computers in Human Behavior*, 8, 249-257.
- Park, E., Mohan, R. & Ponnusamy, V. (2009). Facilitating Learning Through Web 2.0 Collaboration - Cases Across the Culture. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (p. 3764). Chesapeake, VA: AACE.
- Pfeil, U., Zaphiris, P., & Ang, C. (2006) Cultural Differences in Collaborative Authoring of Wikipedia. *Journal of Computer Mediated Communication*, 12, 88-113.
- Safran, C., Helic, C., & Gutl, C. (2007). E-Learning practices and Web 2.0. *Proceedings of ICL 2007* (pp. 1-8). Villach, Austria.
- Selwyn, N. (2007). Web 2.0 applications as alternative environments for informal learning – a critical review. *Paper for OCED-KERIS expert meeting*. Session 6 – alternative learning environments in practice: using ICT to change impact and outcomes.
- Shin, D. & Kim, W. (2008), Applying the technology acceptance model and flow theory to Cyworld user behavior, *CyberPsychology and Behavior*, 11(3), 378-82.
- Shin, D. (2010). Analysis of online social networks: a cross-national study. *Online Information Review*, 34(3), 473-495.
- Singh, N., Zhao, H., & Hu, X. (2003). Cultural adaptation on the web: A study of American companies' domestic and Chinese websites. *International Journal of Global Information Management*, 11(3), 63-80.
- Triandis, H. C. (1988). Collectivism vs. individualism: A reconceptualization of a basic concept in cross-cultural psychology. In C. Verma & C. Bagley (Eds.), *Cross-cultural studies of personality, attitude and cognition*, London: Macmillan, 60-95.
- Triandis, H.C. (1994). *Culture and social behavior*. New York: McGraw-Hill.
- Weller, M., Pegler, C., & Mason, R. (2005). Use of innovative technologies on an e-learning course. *Internet and Higher Education*, 8(1), 61-71.
- Uzuner, S. (2009). Questions of Culture in Distance Learning: A Research Review. *The international Review of Research in Open and Distance Learning*, 10(3), retrieved May 1, 2011 from <http://www.irrodl.org/index.php/irrodl/article/view/690>.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-78.

Factors Affecting Faculty Web Portal Usability

Rex P. Bringula and Roselle S. Basa

College of Computer Studies and Systems, University of the East-Manila, Philippines // rex_bringula@yahoo.com // roselle_basa@yahoo.com

ABSTRACT

The study investigated the factors that might significantly affect web portal usability. Results of the study were intended to serve as inputs for faculty web portal development of the University of the East-Manila. Descriptive statistics utilized questionnaire data from 82 faculty members. The data showed that most of the respondents were relatively young, were Master's degree holders, were skilled in using the computer and the internet, had internet access at home, and were committed to using the web portal. Respondents perceived that the different web portal design-related factors were moderately evident in the existing faculty web portal. Multiple regression analysis showed that information content as a web portal design-related factor was the only significant predictor of web portal usability. Thus, the null hypothesis stating that faculty-related and web portal design-related factors do not significantly affect faculty web portal usability is accepted except Information Content. The study also discussed the guidelines for web portal developers, as well as the limitations and implications of the study for future research.

Keywords

Information content, Web portal design, Web portal usability, UE portal, Usability

Introduction

One of the features of the web is a web portal. A portal, in general, is a gateway to information and services from multiple sources (Tatnall, 2005 cited in Manouselis et al., 2009) that facilitate users' access to the content in one or more learning repositories (Holden, 2003 cited in Manouselis et al., 2009). It has a database that helps users store, locate, and retrieve learning resources (Holden, 2003 cited in Manouselis et al., 2009) in an easy centralized access to all relevant network content and applications (Tatnall & Davey, 2007 cited in Tatnall, 2009). Since the introduction of web portal concept from search engine sites such as Yahoo!, Excite, and Lycos (Tatnall, 2005 cited in Tatnall, 2009), "portals have now become an extremely important aspect of the Web and are now quite ubiquitous" (Tatnall, 2009, p. 1). In fact, Bricolo et al. (2007) found out that web portals are the most visited site among Italian families.

As technology becomes a key tool for good teaching (Adonis, 2006), academic institutions promote and encourage optimization of the Internet technology for information dissemination. The web is becoming an educational medium for universities (Manouselis et al., 2009). Higher education institutions have developed their own web portals because "universities wish to project the impression that they offer the most convenient service and excel in the field of Information Technology—this allows them to attract superior students" (Lee et al., 2009, p. 2). Manouselis et al. (2009) called it educational web portal. This serves as a gateway to information and services of some learning or teaching relevance (Manouselis et al., 2009).

The University of the East (UE) in Manila is one of the universities that recognize the power of the Internet. It continuously uses its financial resources to "connect" to the rest of the world and to support online learning. To realize these goals, UE developed its website and faculty web portal (Figures 1 and 2). In the convenience of their own homes, students and teachers can communicate undisrupted (Figure 2-f). Teachers can post lectures, notes, seat works, assignments, and even quizzes in advance (Figure 2-e).

Students can download these files and be kept informed of the lessons. Teachers can view their academic profile (Figure 2-a), know their teaching load for the current semester (Figure 2-b), and download class lists to verify students' inclusion in the class (Figure 2-c). They can research online through its e-journals and other scholarly works (Figure 2-d). They can post the grades of the students so that they will know their class standing (Figure 2-e). Hence, a faculty web has been an access point for the diverse faculty-related undertakings that include reviewing literature, faculty-student instructional transactions, blogging, student-teacher evaluation, and record keeping.



Figure 1. The UE Faculty Web Portal Log-in Form

Development and maintenance of a web portal is expensive and time-consuming. Therefore, it is imperative that the web portal be designed accordingly to achieve its optimum use. The faculty portal section in the UE website was designed with elements for academic transactions (class list downloads, lecture uploads, student-teacher forums, and evaluation results) and online research (e-journals, e-books, articles, and case studies). However, despite the elements supplied in the design, there is still a need to study the factors that significantly affect the web portal usability. There is a wealth of literature discussing website usability (e.g., Palmer, 2002; Tarafdar & Zhang, 2005; Cappel & Huang, 2007; Pearson & Pearson, 2008; Pang et al., 2009) but very few studies have been conducted to investigate web portal usability.

Figure 2. UE Faculty Web Portal Main Page

This study was conceived to address this gap. Specifically, it aimed to answer the following questions. (1) What are the faculty-related factors in terms of age, highest educational attainment, level of computer applications and Internet skills, commitment of use of faculty web portal, and presence of internet access at home? (2) How evident are the web portal design-related factors, such as ease of navigation, information content, availability, speed, and aesthetics, in the existing faculty web portal? (3) What is the level of usability of the web portal elements (academic

transactions and online research) as to frequency of use? and (4) Do the faculty-related factors and web-portal design-related factors, singly or in combination, affect the faculty web portal usability?

Literature review

Different studies have used the term *usability* in varied ways, thus making it a very confusing concept (Green & Pearson, 2006). There is also no common agreement on a definition of usability (Sindhuja & Dastidar, 2009). Nevertheless, Pressman (2001) attempted to define usability as an attempt to quantify user-friendliness yielding measurable attributes of the users, like (a) skills, both physical and/or intellectual, necessary to learn the system, (b) time required to become efficient in the use of the system, (c) increase in productivity by users who are efficient in utilizing the system, and (d) subjective evaluation of users' attitude towards the system. Usability can also be defined as "the measure of the quality of a user's experience when interacting with a product or system — whether a Website, a software application, mobile technology, or any user-operated device" (Usability.gov, n.d.).

Usability is an important concept in systems development because it is equated to financial gain or loss. Systems with poor usability due to poor website design result in negative financial impacts (Tarafdar & Zhang, 2005). If a website is not usable many users will simply access another site that can meet their needs (Cappel & Huang, 2007). Customers most probably switch to the competitor's sites (Cappel & Huang, 2007).

Landauer (1996 cited in Tarafdar & Zhang, 2005) said that the U. S. economy loses \$30 billion every year in lost productivity due to insufficient use of usability engineering methods in system and website development. Poor website usability costs e-retailers approximately \$4 billion loss in revenue, since many prospective customers do not complete the transactions that they have started. This is a big financial loss to organizations since they are investing millions of dollars in internal and external websites (Pearson & Pearson, 2008).

On the other hand, a usable website that supports customers is associated with higher firm performance (Saeed et al., 2002/3 cited in Pang et al., 2009), increased sales, higher levels of website traffic, improved user performance, increased use of specific features (Nielsen, 2000 cited in Pearson & Pearson, 2008), positive attitude toward online stores, increased stickiness and revisit rates, and stimulated online purchases (Green & Pearson, 2006). However, "despite advancements in technology and significant investments, it is commonly observed that web services implementing business processes have low usability" (Geczy et al., 2011, p. 131). This can be attributed from misalignment between natural characteristics of human interactions in the digital environments and their design and implementation (Geczy et al., 2011).

Numerous research articles attempted to address this problem. Preece et al. (2002) advocated the following design principles to achieve high usability: (a) visibility of system status, match between system and real world, user control and freedom, (b) consistency and standards, (c) user support to recognize, diagnose and recover from errors, (d) error prevention, (e) recognition rather than recall, (f) flexibility and efficiency of use, (g) aesthetic and minimalist design, and (h) help and documentation.

Pearson & Pearson (2008) found out that ease of navigation is a critical component of website usability. Ease of navigation refers to "ease of finding what one desires and knowing where one is in the website" (Sindhuja & Dastidar, 2009, p. 58). It influences website usability (Tarafdar & Zhang, 2005). Ruffini (2001), Palmer (2002), Becker (2005), Seethamraju (2006), Tarafdar & Zhang (2007), Sindhuja & Dastidar (2009), and Zaharias & Poylymenakou (2009) also support this factor. Thus, websites can be improved with regard to navigation (Cappel & Huang, 2007).

Palmer (2002) found out that information content is associated with successful websites. Sindhuja & Dastidar (2009) also found out that information content has significance in the web context because a useful website gives a business-to-consumer (B2C) company a competitive advantage (Pearson & Pearson, 2008). Shneiderman (2005 cited in Pearson & Pearson, 2008) emphasized that website designers must go beyond user friendliness. Website designers should understand the needs of the users. The content of the website should deliver quality information to the users (Cooke, 2003). In other words, designers should look into the relevance (Pearson & Pearson, 2008), depth and breadth (Ruffini, 2001), accuracy (Seethamraju, 2006; Sindhuja & Dastidar, 2009), concurrency (Seethamraju,

2006), and consistency (Sindhuja & Dastidar, 2009) of the information of a website. These ensure the good performance (Tarafdar & Zhang, 2007) and high usability (Tarafdar & Zhang, 2005) of a website.

A regular and consistently available website is one of the factors of a successful website (Tarafdar & Zhang, 2005). If a website is not available, then it cannot be used. If it cannot be used, then customers cannot conduct a transaction (Pearson & Pearson, 2008). Failure to transact greatly affects the performance of a website (Tarafdar & Zhang, 2007). It can be noted that customers perceive the quality of a web by its availability (Seethamraju, 2006). Thus, availability also influences the usability of a website (Tarafdar & Zhang, 2005).

Users feel anxious and lose satisfaction with the website if there is a prolonged delay to access online information (Tarafdar and Zhang, 2005). Nielsen (2000 cited in Pearson and Pearson, 2008) advocated that speed should be considered in designing a website since users' attention diminishes if they will be waiting for more than 10 seconds. Thus, it was perceived as a factor of a quality website (Seethamraju, 2006) and it came out as a predictor of website usability (Tarafdar & Zhang, 2005).

Aesthetics is one of the design criteria in assessing a product (Flowers, 2005) and no business wants to be associated with dull and uninspiring designs (Blunden, 2003). Ruffini (2001) argued that the design of information, selection of graphics, and visual elements (that include color, text, and graphics) are directly related to the intended user. Cyr et al. (2009) found out that aesthetics was considered in a website design. Similarly, Brady & Phillips (2003) found out that an aesthetically pleasing site ranked the highest when the subject users were asked to predict which would be the easiest to use. It is interesting to note that majority of the subject users commented that color was a factor in ranking aesthetic appeal.

Moreover, some studies showed that web aesthetics had positive effects on activation of search on a website and on the intention to purchase on an online store (Wang et al., 2010). Banati et al. (2006) also found out that aesthetics contributed to improving the usability of the site. Becker (2005), Sindhuja & Dastidar (2009), and Zaharias & Poylymenakou (2009) also used this construct.

Different studies also showed that user characteristics can influence the usability of a website. Banati et al. (2005) argued that age could affect the usability of a website. For instance, Milne et al. (2005) said that older people were at risk of being digitally excluded since web content was inappropriately designed with age-related impairments. Milne et al. (2005) further said that the contents of the website were commonly aimed for youth who were technologically adept and familiar with Internet conventions. Mead et al. (2000) also disclosed that older computer users had lower library database search performance than younger adults with similar computer experience. In addition, Becker (2005) revealed that older people were encountering usability problems on e-government sites.

Milne et al. (2005, p. 560) attributed this to age since "the cluttering of small, difficult-to-identify items can make the fairly ubiquitous task of 'pointing and clicking' excessively demanding for many older people; that is, precise manual dexterity becomes more difficult with advanced age, particularly when motion is indirectly mapped from the mouse onto the screen." Sun et al. (2005) confirmed this as they found out that age was a significant predictor of Internet use. More precisely, increased age was associated with less website experience (Laberge & Scialfa, 2005).

Milne et al. (2005) also said that uneducated users were at risk of being excluded as Internet users. As a result, website developers are advised to know the characteristics of the target audience such as their educational background (Ruffini, 2001). Meanwhile, effective design is not absolute since it depends on the skills of individual users (Sandvig & Bajwa, 2004). It is also important to know the user's skills and experiences in using a computer and the Internet. Zhang et al. (2006) found out that the user's skills and experiences were correlated with the user's e-service satisfaction. They also affected the user satisfaction and the intention to use a website. Computer self-efficacy, as Pearson & Pearson (2008) called it, is the ability of an individual to use a computer. It is said that skilled users can perform the necessary work more effectively (Lee, 2008 cited in Lee et al., 2009).

Commitment to a website is the user's intention to continue using the website in the future (Casalo et al., 2007). Hence, a skilled user who is not committed to use a website may result into low website usability. Conversely, a user who is committed to use a website is found to stick (repetitive visits to and use of a preferred website consistently in the future) in using a website (Li et al., 2006).

Finally, in the study of Bricolo et al. (2007) on the home Internet usage among the Italian families, web portals were the most visited site. Rieh (2004, p. 743) explained this finding since “as more people gain at-home access to the Internet, information seeking on the Web has become embedded in everyday life.” Thus, Internet access at home can be an influencing factor of website usability.

Research Paradigm

The foregoing literature review served as basis in formulating the research paradigm shown in Figure 3. The study proposed that the web portal usability could be influenced by (1) faculty-related factors and (2) web portal design-related factors. These factors were regressed, singly and in combination, with the dependent variable (Faculty Web Portal Usability) to determine whether the independent variables could affect the usability of the web portal.

Respondents were asked to rate their level of computer applications and Internet skills. These skills were needed to use the web portal successfully. There were 20 questions on this construct. Respondents were also asked to rate their commitment to the use (i.e., their long-term willingness in using the web portal) of the faculty web portal. The study also aimed at finding out whether respondents had Internet access at home.

Web portal design-related factors made up the second set of constructs. These were subdivided further into five (5) categories with the following definitions.

- Ease of navigation refers to the web portal design characteristic that ensures user's clear understanding of the web portal structure. It allows the user to move easily from section of the portal to another.
- Information content is the web portal attribute pertaining to the relevancy, amount, and usefulness of information found in the web portal. It also refers to the mechanisms on how contents are disseminated.
- Availability depicts the duration of time as dictated by the system administrator or website owner that determines when a web portal is available to the user for access.
- Speed is the response time of the web server of the web portal on the different user's activities.
- Aesthetics is the general look of the web portal. It includes various design attributes such as color, web pages layout, font styles and sizes, and general appearance.

Moreover, the independent variable (Faculty Web Portal Usability) was measured in terms of frequency of use (e.g., Palmer, 2002). It was divided further into academic transactions and online research.

Hypothesis

In accordance with the research paradigm (Figure 3), the null hypothesis given below was tested using statistical tools.

$H_0:$ Faculty-related and Web portal design-related factors do not significantly affect faculty web portal usability.

Data and Methodology

Research Design, Locale, Subjects, Sampling Design and Procedure

This descriptive study was conducted using self-administered questionnaire on faculty members who were selected through random sampling. Using the Slovin's formula with a sampling error computed at 0.10, the minimum sample size of 80 was derived from the 400 faculty members employed at the University of the East-Manila during the First Semester of School Year 2007-2008. Sample size was increased by 37.5% to accommodate potential low completion rate.

Letters of request were sent to the five colleges of the University. All colleges participated in the survey. Twenty (20) survey questionnaires were distributed to each of the six (6) colleges of UE-Manila except College of Law with ten (10). Eighty-two (82) forms were retrieved and these were all used.

Independent Variables

- Faculty-Related Factors**
 - Age
 - Highest Educational Attainment
 - Level of Computer Applications and Internet Skills
 - Commitment to the Use of Faculty Web Portal
 - Presence of Internet Access at home

- Web Portal Design-Related Factors**
 - Ease of Navigation
 - Information Content
 - Availability
 - Speed
 - Aesthetics

Dependent Variable

- Faculty Web Portal Usability**
- Frequency of use**
 - a. academic transactions
 - class list download
 - lecture upload
 - student-teacher forums
 - evaluation results
 - b. online research
 - e-journals
 - books
 - articles
 - case studies

Figure 3. The Research Paradigm Showing the Factors That Might Affect the Faculty Web Portal Usability

Research Instrument

Three experts validated the content of the questionnaire. Questionnaire pre-testing was conducted on 30 faculty members of the College of Computer Studies and Systems. They were selected in the questionnaire pre-testing because of their familiarity with the UE web portal. Modifications (e.g., inclusion and/or exclusion of confusing, leading, or irrelevant questions) were done based on three experts' recommendations and on the results of pre-testing. After these modifications, the questionnaires were distributed to the faculty members of the six colleges of the University.

Faculty-related factors were composed of age, educational attainment, availability of internet access at home, level of computer applications and internet skills and commitment to the use of faculty web portal. Levels of computer applications and internet skills were measured using a 5-point scale ranging from "highly skilled" (5) to "not skilled" (1) (See Table 1) while commitment to the use of the faculty web portal was measured using a 5-point scale ranging from "highly committed" (5) to "not committed" (1) (See Table 1).

(a)

(b)

Figure 4. Sample Webpage Facilitating (a) Academic Transaction and (b) Online Research

The web portal design-related factors, which dwelt on the perception of the respondents on how the features of the portal were appropriately designed for academic transactions (Figure 4-a) and online research (Figure 4-b), were likewise evaluated using a 5-point scale ranging from “highly evident” (5) to “not evident” (1) (See Table 1).

The frequency of weekly access was measured through a 5-point scale ranging from “very often” (5) to “never” (1) (See Table 1).

Table 1. The 5-Point Scale and Its Verbal Interpretation and Mean Range

| Verbal Interpretation | Weight | Mean Range |
|--|--------|-------------|
| Highly (skilled/committed/evident) / Very Often | 5 | 4.51 – 5.00 |
| Skilled/Committed/Evident / Often | 4 | 3.51 – 4.50 |
| Moderately (skilled/committed/evident) / Sometimes | 3 | 2.51 – 3.50 |
| Slightly (skilled/committed/evident) / Seldom | 2 | 1.51 – 2.50 |
| Not (skilled/committed/evident) / Never | 1 | 1.00 – 1.50 |

Factor analysis and Cronbach’s alpha was used to determine the validity and reliability of the constructs, respectively. All constructs (See Table 2) were found to be valid (factor loading greater than 0.50) and reliable (above the minimum criterion of 0.70) (Pallant, 2001; George & Mallery, 2009).

Table 2. Validity and Reliability of Constructs

| Constructs | Number of questions | Factor loadings | Cronbach’s alpha |
|-----------------------|---------------------|-----------------|------------------|
| Ease of Navigation | 5 | 0.814 | 0.967 |
| Information Content | 5 | 0.883 | 0.891 |
| Availability | 5 | 0.847 | 0.923 |
| Speed | 5 | 0.838 | 0.965 |
| Aesthetics | 5 | 0.840 | 0.963 |
| Academic Transactions | 4 | 0.931 | 0.872 |
| Online Research | 4 | 0.931 | 0.963 |

Frequency counts, percentages, and mean were used to describe the respondents and their responses statistically. Multiple regression analysis at 1% level of probability and 99% reliability was used to determine which of the faculty- and web portal design-related factors would significantly affect the usability of faculty web portal.

Results, Analysis and Evaluation

The Faculty-Related Factors

As seen in Table 3, most of the respondents were relatively young, were Master’s degree holders, were committed to the use of the faculty web portal, and had Internet access at home. It also shows that respondents were skilled in using different computer applications and the Internet (e.g., starting up, rebooting, and shutting down a computer, creating/deleting/renaming a file, downloading and uploading files, searching e-journals, etc.). The alpha value of this construct is 0.985.

Table 3. Summary of Findings on Faculty-Related Factors

| Faculty-Related Factors | Findings |
|--|---|
| Age | Most of the respondents were in the age bracket of 30–39 ($f=34$, 41%). |
| Highest Educational Attainment | Most of the respondents attained master’s degree ($f=54$, 66%). |
| Level of Computer Applications and Internet Skills | Respondents are <i>skilled</i> in using different computer applications and the Internet (mean = 4.14). |
| Commitment to the Use of Faculty Web Portal | Respondents were committed (mean = 3.61) to the use of the faculty web portal. |
| Presence of Internet Access at Home | Sixty-six percent ($f=54$, 66%) of the respondents had internet access at home. |

Meanwhile, as shown in Table 4, Availability got the lowest mean rating of 3.01 (moderately evident). Aesthetics got the highest mean rating of 3.73 (evident). The overall perception of the respondents yielded to 3.30 (moderately evident). This shows that faculty web portal with respect to design, content structure, and organization is moderately acceptable to the respondents.

However, respondents perceived that the design of the faculty web portal in accordance with the web the different web portal-related factors as moderately evident (grand mean = 3.30). According to Tarafdar & Zhang (2007), design factors, such as ease of navigation, information content, speed, and availability are characteristics of a website that highly influence usability. Unfortunately, these are perceived to be moderately evident. Follow-up questions showed that this is mainly because (a) respondents encounter dead links – links that do not function while surfing the web portal, (b) the portal moderately support downloading and uploading mechanisms of seat works, assignments, quizzes, etc., (c) electronic copies of research, thesis, and dissertations are not available online, and (d) no chatting facility.

Table 4. Mean Level of Consideration of the Respondents on the Faculty Web Portal as to Different Web Portal Design-Related Factors

| Web Portal Design-Related Factors | Mean | Verbal Interpretation |
|-----------------------------------|-------------|---------------------------|
| Ease of Navigation | 3.32 | Moderately Evident |
| Information Content | 3.35 | Moderately Evident |
| Availability | 3.01 | Moderately Evident |
| Speed | 3.10 | Moderately Evident |
| Aesthetics | 3.73 | Evident |
| Grand Mean | 3.30 | Moderately Evident |

Table 5 shows that faculty respondents only use the faculty web portal from time to time (grand mean = 2.74, Sometimes).

Table 5. Frequency of Use as to Academic Transactions

| Academic Transactions | Mean | Verbal Interpretation |
|--------------------------------|-------------|-----------------------|
| Downloading of Class List | 3.00 | Sometimes |
| Uploading of Lecture Materials | 2.57 | Sometimes |
| Student-Teacher Forums | 2.39 | Seldom |
| Evaluation Results | 3.00 | Sometimes |
| Grand Mean | 2.74 | Sometimes |

Student-Teacher Forums yielded the lowest score of 2.39 (Seldom) while Downloading of Class List and Viewing of Evaluation Results got the highest scores of 3.00 (Sometimes). Faculty respondents use the faculty web portal only from time to time (grand mean = 2.74).

These findings can be attributed to the following reasons. (1) Respective Colleges give class lists to the faculty members and the faculty do not see the need to download the class lists, or if they do download the class lists, they just need to acquire a softcopy once. (2) Student evaluation is conducted once every semester. Therefore, viewing its results could happen at the later part of the semester. (3) Uploading of files through the web portal is only limited to one megabyte and limited only to a specific type of files (e.g., spreadsheet and PDF documents cannot be uploaded). (4) The nature of the present items of the portal does not require to be accessed on a daily basis, e.g., viewing of evaluation results and downloading of class lists. (5) There is lack of specific policy regarding the use of the portal. It can be noted that there is no University-wide existing policy regarding the use of faculty web portal.

Meanwhile, from Table 6, all online research materials were only used sometimes (grand mean = 2.72). The low usability of online research cannot be attributed to traditional research in the library. Library usage statistics showed that during the school years 2005-2006, 2006-2007, and 2007-2008, the daily average number of faculty who visits the library was only 15, 12, and 12, respectively.

Table 6. Frequency of Use as to Online Research

| Online Transactions | Mean | Verbal Interpretation |
|---------------------|-------------|-----------------------|
| Electronic Journals | 2.71 | Sometimes |
| Books | 2.70 | Sometimes |
| Articles | 2.78 | Sometimes |
| Case Studies | 2.70 | Sometimes |
| Grand Mean | 2.72 | Sometimes |

It can be noted that even though faculty-respondents are skilled in performing various computer- and Internet-related activities, have a high sense of commitment to the use of the faculty web portal, and have internet connection at home (Table 3), the faculty web portal is still not used frequently. The faculty web portal log shows that out of 695 faculty members, only 12 (2%) logged on to the portal for January, 2008, 19 (3%) logged on February, 39 (6%) logged on March, 87 (12%) logged on June, 120 (17%) logged on July, and 55 (8%) for the first week of August.

Factors That Affect Faculty Web Portal Usability

Table 8 shows the stepwise regression of faculty web portal usability on faculty-related and web portal design-related factors. Only information content was found to be a significant predictor ($\beta = 0.57$; $p\text{-value} = 0.000$) of faculty web portal usability. The adjusted R^2 shows that 31% (Adjusted $R^2 = 0.31$) in the variation in usability of the faculty web portal was accounted to information content. The prediction was unlikely to have arisen from sampling error ($F\text{-value} = 37.66$, $p\text{-value} < 0.01$).

Table 8. Regression of Faculty Web Portal Usability on Faculty-Related and Web Portal Design-Related Factors

| Predictor | Adjusted R ² | F-value | Beta | p-value |
|-----------------------------|-------------------------|---------|------|---------|
| Information Content* | 0.31 | 37.66 | 0.57 | 0.000 |

* $p\text{-value} < 0.01$ level of significance

The result agreed with the study of Tarafdar & Zhang (2005) that information content influences web usability. It also supported the studies of Ruffini (2001), Palmer (2002), Seethamraju (2006), Tarafdar & Zhang (2007), Pearson & Pearson (2008), and Sindhuja & Dastidar (2009). It can be noted that information content refers to the website attribute pertaining to the variety, amount, quality, and usefulness of information found in the website, as well as various mechanisms on how information may be conveyed from one type of use to another. This is the main objective of the faculty web portal. Ensuring the comprehensiveness of the content of the web portal will help further realize the objectives of the web portal. This implies that in evaluating the effectiveness of a web portal, the questionnaire should surface the rich information content of a web portal.

Implications for Web Portal Developers

The finding shown in Table 8 implies that web portal programmers should develop a web portal based on the needs of the users. During the development phase of the web portal, direct users should be involved in the development to ensure that the elements in the portal are really needed. Banati et al. (2005) said that different stakeholders influenced website usability. Among these stakeholders are the actual users.

Users are viewed as important sources of knowledge (Axtell et al., 1997 cited in Sugar, 2001) in the development of the system. Thus, actual interviews, not intellectual guesses, should be deployed to direct users to ascertain their needs. This participatory design in the development of computer systems needs full participation from the users (Greenbaum & Kyng, 1991 cited in Sugar, 2001) since it can avoid confusion or irritation of the user (Sandvig & Bajwa, 2004). Knowing the target users (Ruffini, 2001; Flowers, 2005) and their needs can be done through focus group interview. This is found to be a very efficient and effective approach in designing a website (Choe et al., 2006).

For examples, Abdous (2005), in an attempt to achieve higher educational impact of the web portal, included faculty-centered tools such as syllabus generator and curriculum mapping tool. Gallant et al. (2007) employed participatory

design methods in the development of hospital websites. Wegener et al. (2004) designed a website for a library by putting themselves in their students' shoes. In this manner, if web developers supplied contents relevant to the users during the portal development, then it can be expected that the web portal will be highly usable.

Conclusion and Recommendations

The results of the study show that among all the factors studied, only information content of the web portal design-related factors significantly affected faculty web portal usability. Thus, the null hypothesis stating that faculty-related and web portal design-related factors do not significantly affect faculty web portal usability is accepted except Information Content.

This study is considered to be a pioneering study in web portal usability in the Philippines. However, there are lots of gap in the literature to fill in. In the light of the limitations, findings, and conclusion presented, the following recommendations are set forth.

- The information content of faculty web portal can be enhanced. Include chatting capabilities, faculty e-mail query, researchers' hub, or online book reservations. Involve faculty members in the development of the web portal. In other words, know the target users (Ruffini, 2001).
- Factors other than those studied here can be included such as ease of use (Hart et al., 2008; Ryan & Rao, 2008; Aljukhadar & Senecal, 2009), retention (Rodgers & Negash, 2007) and attitude (Aljukhadar & Senecal, 2009) towards using the web portal. Since humans dealt with a computer system, human factors, such as those stated earlier, might have effect on web portal usability. Also, instead of self-rated frequency of use, actual number of hours of use as indicator of frequency of use can also be utilized as the independent variable.
- Further investigation can be conducted to find out the causes of low usability of faculty web portal usability in spite of committed, internet-connected, and skilled faculty members. Thus, in accordance with the previous recommendation, qualitative-quantitative techniques can be utilized to investigate this.
- Further studies can investigate the usability of the web portal in two poles: one at the side of the faculty and the other at the side of the students. Usability of the two web portals can only be realized if the elements of the two web portals match.
- Though the study found out that respondents were not novice computer and Internet users, it is not clear whether they were novice web portal users. User experience categorized as novice, expert and professional can influence website usability (Banati et al., 2005).
- Proper training or orientation on the use of the faculty web portal should be conducted. Formulate the institutional policy on the use of the web portal. This could address the issue of novice users.
- User satisfaction as a measure of usability can also be investigated. This is necessary since user satisfaction guides the viewpoint of a user towards a website and hence its usage (Banati et al., 2006). Palmer (2002) and Abdinnour-Helm et al. (2005) used satisfaction as a measure of the performance of a website.
- This study did not deal on the usefulness and effectiveness of the web portal. Hence, usefulness of the portal based on its purpose (Ruffini, 2001) and its effectiveness (Tatnall, 2009) can be conducted.
- Lastly, an investigation of its reasons of adoption (or non-adoption), advantages and problems associated with its use (Tatnall, 2009) can also be initiated.

Acknowledgments

The authors are greatly indebted to Dr. Panfilo O. Domingo, Dr. Ester A. Garcia, Dr. Olivia C. Caoili, Dean Constantino T. Yap, Dean Rodany A. Merida, Felicito C. Nuarin, Dr. Socorro R. Villamejor, and to the two anonymous reviewers.

References

- Abdinnour-Helm, S. F., Chaparro, B. S., & Farmer, S. M. (2005). Using the end-user computing satisfaction (EUCS) instrument to measure satisfaction with a web site. *Decision Sciences*, 36(2), 341–364.

- Abdous, M. (2005). Designing a faculty portal framework and architecture. In *21st Annual Conference on Distance Teaching and Learning*. Retrieved October 17, 2008, from <http://www.uwex.edu/disted/conference>.
- Adonis, L. A. (2006). Technology in schools. *Manager: British Journal of Administrative Management*, 52, 14–15.
- Aljukhadar, M., & Senecal, S. (2009). How the website usability elements impact performance. *Value Creation in E-Business Management*, 36(2), 113–130.
- Banati, H., Bedi, P., & Grover, P.S. (2006). Evaluating web usability from the user's perspective. *Journal of Computer Science*, 2(4), 314–316.
- Becker, S.A. (2005). E-government usability for older adults. *Communications of the ACM*, 48(2), 102–104.
- Blunden, P. (2003). Why usability matters. *New Media Age, Centaur Communications Limited*. Retrieved December 19, 2010, from <http://www.highbeam.com/doc/1G1-105085336.html>.
- Brady, L., & Phillips, C. (2003). *Aesthetics and usability: A look at color and balance*. *Usability News*, 5(1). Retrieved October 17, 2007, from <http://www.surl.org/usabilitynews/51/pdf/Usability%20News%2051%20-%20Brady.pdf>.
- Bricolo, F., Gentile, D. A., Smelser, R. L., & Serpelloni, G. (2007). Use of the computer and Internet among Italian families: First national study. *CyberPsychology & Behavior*, 10(6), 789–797.
- Cappel, J., & Huang, Z. (2007). A usability analysis of company websites. *The Journal of Computer Information Systems*, 48(1), 117–123.
- Casalo, L., Flavia'n, C., & Guinaliu, M. (2007). The influence of satisfaction, perceived reputation and trust on a consumer's commitment to a website. *Journal of Marketing Communications*, 13(1), 1–17.
- Choe, P., Kim, C., Lehto, M. R., Lehto, X., & Allebach, J. (2006). Evaluating and improving a self-help technical support web site: Use of focus group interviews. *International Journal of Human-Computer Interaction*, 21(33), 333–354.
- Cooke, M. (2003). Why real people are vital to good usability. (Analyst Speak). *New Media Age, Centaur Communications Limited*. Retrieved February 15, 2011, from <http://www.highbeam.com/doc/1G1-105366024.html>.
- Cyr, D., Head, M., Larios, H., & Pan, B. (2009). Exploring human images in website design: A multi-method approach. *MIS Quarterly*, 33(3), 539–566.
- Flowers, J. (2005). Usability testing in technology education. *The Technology Teacher, International Technology Education Association*. Retrieved December 19, 2010, from <http://www.highbeam.com/doc/1G1-133139184.html>.
- Gallant, L., Irizarry, C., & Kreps, G. L. (2007). User-centric hospital web sites: A case of trust and personalization. *e-Service Journal*, 5(2), 5–26.
- Geczy, P., Izumi, N., & Hasida, K. (2011). Foundations for effective portal service management. *Global Journal of Business Research*, 5(2), 131–141.
- George, D., & Mallory, P. (2009). *SPSS for Windows step by step: A simple guide and reference 16.0 update* (9th Ed.). Boston: Pearson Education.
- Green, D., & Pearson, J. M. (2006). Development of a web site usability instrument based on ISO 9241-11. *The Journal of Computer Information System*, 47(1), 66–72.
- Hart, T. A., Chaparro, B. S., & Halcomb, C. G. (2008). Evaluating websites for older adults: Adherence to 'senior-friendly' guidelines and end-user performance. *Behaviour & Information Technology*, 27(3), 191–199.
- Laberge, J. C., & Scialfa, C. T. (2005). Predictors of web navigation performance in a life span sample of adults. *Human Factors*, 47(2), 289–302.
- Lee, H. S., Choi, Y. H., & Jo, N. O. (2009). Determinants affecting user satisfaction with campus portal services in Korea. *Journal of Internet Banking and Commerce*, 14(1), 1–18.
- Li, D., Browne, G. & Wetherbe, J. (2006). Why do internet users stick with a specific web site? A relationship perspective. *International Journal of Electronic Commerce*, 10(4), 105–141.
- Manouselis, N., Kastrantas, K., Sanchez-Alonso, S., Caceres, J., Ebner, H., Palmer, M., & Naeve, A. (2009). Architecture of the Organic.Edunet web portal. *International Journal of Web Portals*, 1(1), 71–91.
- Mead, S. E., Sit, R. A., Rogers, W. A., Jamieson, B. A., & Rousseau, G. K. (2000). Influences of general computer experience and age on library database search performance. *Behavior & Information Technology*, 19(2), 107–123.
- Milne, S., Dickinson, A., Carmichael, A., Sloan, D., Eisma, R., & Gregor, P. (2005). Are guidelines enough? An introduction to designing web sites accessible to older people, *IBM Systems Journal*, 44(3), 557–571.
- Pallant J. (2001). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows version 10*. Buckingham: Open University Press.
- Palmer, J. W. (2002). Web site usability, design, and performance metrics. *Information Systems Research*, 13(2), 151–167.
- Pang, M.-S., Suh, W., Kim, J., & Lee, H. (2009). A benchmarking-based requirement analysis methodology for improving web sites. *International journal of Electronic Commerce*, 13(3), 119–162.

- Pearson, J. M., & Pearson, A. (2008). An exploratory study into determining the relative importance of key criteria in web usability: A multi-criteria approach. *The Journal of Computer Information Systems*, 48(4), 115–127.
- Preece, J., Rogers, Y., & Sharp M. (2002). *Interaction design, beyond human – computer interaction*. New York: John Wiley & Sons.
- Pressman, R. (2001). *Software engineering: A practitioner's approach* (5th Ed.). Boston: McGraw-Hill.
- Rieh, S. Y. (2004). On the web at home: Information seeking and web searching in the home environment. *Journal of the American Society for Information Science and Technology*, 55(8), 743–753.
- Rodgers, W., & Negash, S. (2007). The effects of web-based technologies on knowledge transfer. *Communications of the ACM*, 50(7), 117–122.
- Ruffini, M. F. (2001). Blueprint to develop a great web site. *T H E (Technological Horizons in Education)*. Retrieved October 21, 2009, from <http://thejournal.com/Articles/2001/03/01/Blueprint-to-Develop-a-Great-Web-Site.aspx>.
- Ryan, C., & Rao, U. (2008). Holiday users of the Internet – Ease of use, functionality and novelty. *International Journal of Tourism Research*, 10, 329–339.
- Sandvig, J. C., & Bajwa, D. (2004). Information seeking of university web sites: An exploratory study. *The Journal of Computer Information Systems*, 45(1), 13–22.
- Seethamraju, R. (2006). Web quality – A study of user perceptions. *IIMB Management Review*, 18(1), 15–24.
- Sindhuja, P. N., & Dastidar, S. G. (2009). Impact of the factors influencing website usability on user satisfaction. *The IUP Journal of Management Research*, 8(12), 54–66.
- Sugar, W. A. (2001). What is so good about user-centered design? Documenting the effect of usability sessions on novice software designers. *Journal of Research on Computing in Education*, 33(3), 235–250.
- Sun, P., Unger, J. B., Palmer, P. H., Gallaher, P., Chou, C., Garbanati, L., Sussman, S., & Johnson, C. (2005). Internet accessibility and usage among urban adolescents in Southern California: Implications for web-based health research. *CyberPsychology & Behavior*, 8(5), 441–453.
- Tarafdar, M., & Zhang, J. (2005). Analyzing the influence of web site design parameters on web site usability. *Information Resource Management Journal*, 18(4), 62–80.
- Tarafdar, M., & Zhang, J. (2007). Determinants of reach and loyalty – A study of website performance and implications for website design. *The Journal of Computer Information Systems*, 48(2), 16–24.
- Tatnall, A. (2009). Gateways to portals research. *International Journal of web portals*, 1(1), 1–15.
- Usability.gov. (n.d.). *What does usability measure*. Retrieved January 2, 2010, from <http://www.usability.gov/basics/index.html>.
- Wang, Y. J., Hong, S., & Lou, H. (2010). Beautiful beyond useful? The role of web aesthetics. *The Journal of Computer Information Systems*, 50(3), 121–129.
- Wegener, D. R., Gog-Ong, A.M., & Lim, M.L. (2004). Web usability testing in a polytechnic library. *The Australian Library Journal*, 53(2). Retrieved January 2, 2011, from <http://alia.org.au/publishing/alj/53.2/full.text/wegener.html>.
- Zaharias, P., & Poylymenakou, A. (2009). Developing a usability evaluation method for e-learning applications: Beyond functional usability. *International Journal of Human-Computer Interaction*, 25(1), 75–98.
- Zhang, X., Prybutok, V., & Huang, A. (2006). An empirical study of factors affecting e-service satisfaction. *Human Systems Management*, 25, 279–291.

Appendix

Guidelines for Web Portal Developers

1. Know your intended users. Know what they need, their skills, and their purpose of usage of the web portal.
2. Make an actual interview in a representative sample of the user.
3. Involve users in the development of the web portal. Make a constant interaction with them during the development and deployment phase of the web portal.
4. Ensure that educational materials, such as electronic journals, e-books, and other scholarly manuscripts are available and accessible on the web portal.
5. Ensure a reliable uploading and downloading mechanisms of the web portal. This will enable the faculty to upload academic files for students' perusal.
6. Provide a reasonable memory space for uploading of files.
7. Ensure that the web portal can hold different types of files. It should not only be limited to word or PowerPoint documents.
8. Test the web portal before implementing to a representative sample. Testing the web portal to the actual user will identify the pitfalls of the software.
9. Provide orientation and help. All users will definitely pass through the stage of novice users. An orientation on the purpose, navigation, and usage of the web portal will be very beneficial to attain the web portal high usability.
10. Provide feedback mechanisms. The purpose and the problem of the usage of the web portal together with the emergence of new web technologies will make an existing portal obsolete. A feedback mechanism will greatly help in improving a web portal.
11. Develop a web portal that can log the actual hours of usage of the users. The actual hours of usage of the users are strong indication on the commitment to the use of the web portal. This is very beneficial on measuring the usability of a web portal.

Mining Learning Preferences in Web-based Instruction: Holists vs. Serialists

Natalie Clewley¹, Sherry Y. Chen^{1,2*} and Xiaohui Liu¹

¹School of Information Systems, Computing, and Mathematics, Brunel University, Uxbridge, United Kingdom //

²Graduate Institute of Network Learning Technology, National Central University, Taiwan //

Natalie.Clewley@brunel.ac.uk // Sherry@cl.ncu.edu.tw // Xiaohui.Liu@brunel.ac.uk

*Corresponding author

ABSTRACT

Web-based instruction programs are used by learners with diverse knowledge, skills and needs. These differences determine their preferences for the design of Web-based instruction programs and ultimately influence learners' success in using them. Cognitive style has been found to significantly affect learners' preferences of web-based instruction programs. However, the majority of previous studies focus on Field Dependence/Independence. Pask's Holist/Serialist dimension has conceptual links with Field Dependence/Independence but it is left mostly unstudied. Therefore, this study focuses on identifying how this dimension of cognitive style affects learner preferences of Web-based instruction programs. A data mining approach is used to illustrate the difference in preferences between Holists and Serialists. The findings show that there are clear differences in regard to content presentation and navigation support. A set of design features were then produced to help designers incorporate cognitive styles into the development of Web-based instruction programs to ensure that they can accommodate learners' different preferences.

Keywords

Learning preferences, Holist, Serialist, Pask, Field Dependence, Field Independence

Introduction

The World Wide Web (Web) provides an extremely large and dynamic information resource (Ma, Pant, and Sheng, 2007) and is currently being applied extensively as an important means for information dissemination (Harumoto, et al., 2005). In other words, using the Web has become an essential part of our daily life (Kirkwood, 2008). In particular, Web-based Instruction (WBI) has become increasingly popular in educational settings (Brotherton and Abowd 2004). Due to such popularity, WBI programs are used by many different students, each of which has their own set of personal preferences because they have different backgrounds, knowledge, and skills. In other words, human factors, namely those individual characteristics that can potentially affect the design of human-computer interaction (Sears and Jacko, 2009), are a necessary consideration in the design of WBI programs. There are many human factors, such as cognitive styles (e.g., Chen and Macredie, 2004), gender differences (e.g., Roy and Chi, 2003) and prior knowledge (e.g., Mitchell, Chen, and Macredie, 2005). Among them, cognitive style, which describes and explains differences in the preferred strategies for information representation and processing among individuals (Riding and Rayner, 1998), is particularly widely studied in the area of WBI because it has been shown to have a great effect on learners' preferences (Ford and Chen, 2001). The most widely studied cognitive style dimension is that of Witkin's (1976) Field Dependence/Independence. As showed in previous studies (e.g. Chen and Macredie, 2004; Chen and Liu, 2008), the Field Dependence/Independence dimension has considerable effects on learners' preferences for the design of WBI.

On the other hand, Pask's (1979) Holist/Serialist dimension has conceptual links to the Field Dependent/Independent dimension and such links suggest that this dimension of cognitive style may also play an influential role in the design of WBI programs. As suggested by Ford (2000), the Holist/Serialist dimension has potential for adapting computing-based systems to the needs of each learner. However, previous literature has paid less attention to this dimension of cognitive style. Thus, there is a need to examine how the Holist/Serialist dimension affects learners' preferences for the design of WBI programs. To this end, the study presented in this paper will investigate the influence of this particular cognitive style dimension on learners' preferences for the design of WBI programs.

Furthermore, this study will apply data mining techniques to analyze the relationship between learners' cognitive styles and the influence this has on their preferences for WBI programs because our previous work shows that employing the use of data mining techniques can help to identify relationships that were previously hidden by statistical analyses (e.g. Chen and Liu, 2008). Unlike our previous work, we, however, use multiple data mining techniques. Firstly, two families of classifiers are used to select relevant features and then three types of decision

trees are applied to illustrate learners' preferences. With such an approach, this study will not only contribute to the understanding of how Pask's (1979) Holist/Serialist dimension influences learners' preferences, but also propose a new way to conduct data analyses

In this vein, this paper begins by analyzing previous research relating to WBI and cognitive styles and continues by providing a background on the data mining techniques used in this study. The methodology used to conduct the empirical study is described in section 3, followed by the discussion of the findings in section 4. This section also suggests a set of design points drawn from the findings that should be included in an interface to account for learners' cognitive styles. Conclusions are then presented in section 5 where suggestions for future work are also proposed.

Related Works

Web-Based Instruction

Web-based instruction (WBI) programs provide flexible teaching and learning environments for students through the provision of non-linear learning (Pituch & Lee, 2006), as students have the freedom to control their learning by themselves, for example, through the use of different navigational tools. Due to such freedom, WBI programs are very attractive to students. With the expanding usage of such programs, the ability to effectively match the interface design with the increased diversity in students' preferences becomes vital to their success. Thus, there is a need to examine the influences of human factors, among a variety of which cognitive style has been identified as one of the most pertinent factors because it refers to a person's information processing habits, representing an individual's typical mode of perceiving, thinking, remembering, and problem solving (Messick 1976). It has also been suggested that matching cognitive styles to the design of WBI programs can lead to better learning performance (Ford and Chen, 2001).

In particular, previous studies found that Witkin's (1976) Field Dependence/Independence affects how students react to WBI programs. For example, Lu, Yu, and Liu (2003) found that Field Dependent learners used the teaching notes and class resources more than Field Independent learners. Lee et al (2009) found that Field Independent learners used the back/forward buttons more frequently and spent less time on navigation than their Field Dependent counterparts. Additionally, Field Dependent users were found to use the main menu more often and had more repeated visits. Chen and Liu (2008) found that Field Independent learners preferred to use the alphabetical index whilst Field Dependent learners preferred the hierarchical map. These clear differences permit designers to accommodate these preferences in interface designs, allowing for more effective use of WBI.

Table 1: Differences between Holist and Serialist characteristics (Pask, 1979)

| Holists | Serialists |
|--|---|
| • Take a global approach and create conceptual links between objects early on. | • Take an analytical approach, examining individual topics before forming conceptual links. |
| • Are able to move between theory and real world examples from the beginning. | • Analyzes theory or real world examples separately, only joining together if necessary. |
| • Broad focus; likes to have more than one thing on the go at the same time. | • Narrow focus; prefers to focus on completing one task before moving on to the next. |
| • Internally directed. | • Externally directed. |

In addition to the Field Independence/Dependence cognitive style dimension, another dimension of cognitive style, i.e., Pask's Holist/Serialist (1979), was also found to be an influential factor to student learning. According to Pask's description, Holists will adopt a style termed comprehension learning which involves building descriptions of what is known whereas the Serialists will adopt a style termed operation learning which is concerned with the mastering of procedural details (Pask, 1979). Additionally, Jonassen and Grabowski (1993) describe the Holists as preferring to process information in a 'whole-to-part' sequence. In contrast, Serialists are described as preferring a 'part-to-whole' processing of information. As shown in Ford's study (1993), Holists strongly favored a global picture by using a map. On the other hand, Serialists preferred to use an index to access information. A detailed comparison of the differences between Holists and Serialists preferences can be seen in Table 1.

As shown in this Table, Holists will prefer to adopt a global approach and focus on piecing together a broad conceptual overview before then looking further to fit other details in, whereas Serialists prefer to use a local approach, examining one object at a time and only then concentrating on linking objects to the building of the conceptual map (Pask, 1979). More specifically, the former develop an overall and general understanding of the scope and structure of learning tasks at hand and gradually shift attention to details that fill in the structure while the latter would tackle individual details first, connect the separate topics, and finally form the overall picture (Ford, 2000). In other words, Holists share similar characteristics to those of Field Dependent users, who tend to emphasize on an overall picture of the subject content. Conversely, Serialists are like Field Independent users, who tend to build up procedural understanding step-by-step. In summary, these two dimensions of cognitive style have some similarities, but Pask's Holist/Serialist (1979) is rarely examined in the area of WBI. To address this issue, this dimension of cognitive style is considered in this study. Unlike most of previous studies which used traditional statistics to analyze data, a data mining approach is applied in this study because it can discover hidden patterns (Chen and Liu, 2008).

Data Mining

Although data mining techniques are more traditionally applied in the areas of Bioinformatics or E-Business, there is an increasing interest in applying these techniques in educational settings with the aim of gaining a deeper understanding of students and their learning environments (Romero and Ventura, 2007). Lack of a deep and sufficient understanding in e-learning systems may prevent to deliver high-quality services, but data mining is helpful to bridge this gap (Ayesha, et al., 2010). Thus, some research attempted to use data mining techniques to e-learning systems, including personalizing distance education courses and e-text books (Romero and Ventura, 2007). For instance, Tang et al. (2000) constructed a personalized web tutor tree by using a key-word-driven text mining algorithm to select articles for distance learning students. Additionally, Chen, Li, Wang, and Jia (2005) used web content mining to develop an e-textbook where a ranking strategy was employed to evaluate the web page suitability and to extract concept features and build concept hierarchies.

These works demonstrate that data mining is advantageous to automatically model learners' preferences. However, the analysis of the relationship between cognitive styles and the corresponding user preferences for WBI programs has mainly been investigated through statistical analyses (e.g. Graff, 2003), which require assumptions to be made beforehand, introducing bias if these assumptions are not entirely accurate. In contrast, data mining is discovery driven so there is no need to have any assumptions in advance (Romero and Ventura, 2007). More specifically, data mining is the search for valuable information within large volumes of data (Hand, Mannila, & Smyth, 2001). This valuable information can then be used to predict, model or identify interrelationships (Urtubia et al, 2007) without the need to predefine underlying relationships between dependent and independent variables (Chang and Chen, 2005). However, human preferences data is often 'noisy' and full of inaccurate information, which can lead to biased results. Therefore, a pre-processing step, such as feature selection, is necessary to sift through the data so that only the most relevant subsets are included in the mining process (Bishop, 1995).

Feature selection methods can be divided into two categories, including filters and wrappers. The former uses statistical methods to produce a ranking of features whereas the latter uses classifiers to evaluate small subsets based on the interactions between features (Yang and Olafsson, 2006). The wrappers generally perform better than the filters (Raman and Ioerger, 2002) so the wrappers were taken into account in this study. However, the outcomes of the wrappers depend on the type of classifier used. For example, Bayesian Networks and Nearest Neighbour are two of the most popular families of classifiers. Classifiers from these families have different architectures and therefore have different biases associated with them. For example, Bayesian Network classifiers focus on features that maximize/minimize a scoring metric whereas those from the Nearest Neighbour family focus on features that are deemed the 'closest' by an imposed distance metric (Huan and Lei, 2005). This study addresses attempts to overcome this issue by comparing the accuracy of feature sets obtained using classifiers from various families, instead of choosing just one classifier or one family.

After the data has been pre-processed and contains only the most relevant features, classification, which is a data mining technique that uses algorithms to find models that describe a data class or concept (Han and Kamber, 2006), can be used to model the preferences of different types of learners. One of the most popular classification tools are decision trees, which are employed to discover rules and relationships by systematically dividing information

contained within data (Chen, Hsu and Chou, 2003). Data is classified by constructing tree-like structures through a series of Boolean functions, yes/no questions based on the characteristics of a set of variables, until a pre-defined level is reached. Thus, the hierarchical structure created can help researchers easily understand hidden relationships within the dataset (Lee et al, 2009). Decision trees have been widely used to classify user preferences. For example, Liu and Kešelj (2007) used decision trees to classify users' Web navigational patterns with the aim of predicting which pages were more likely to be visited next.

Due to the success of decision trees in such studies, this study will use decision trees to illustrate the differences in preferences between Holists and Serialists. A review of existing literature shows that the most widely used decision tree algorithms include Classification and Regression Trees (CART, Breiman et al., 1984), C4.5 (Quinlan, 1993) and CN2 (Clark and Nibbet, 1989). A difference amongst these three algorithms is the process of model development. C4.5 algorithms split a tree model into as many results as necessary whereas the CART algorithm can only support binary splits. Unlike CART and C4.5, CN2 is a rule based algorithm that produces an ordered list of IF-THEN rules that can be visually demonstrated on a decision tree. Due to these differences, using a single decision tree may cause some biases. Thus, all of these three algorithms will be applied in this study and then we use the one which produces the highest accuracy to illustrate the different preferences of Holists and Serialists. By doing so, more reliable results can be obtained.

Methodology Design

Empirical Study

The research question examined in this study is how Holists and Serialists show different preferences to the design of a WBI program. The students from a UK university were invited to participate in the study by email. It was indicated in the email that all of the participants were requested to have the basic computing and Internet skills, e.g., the abilities of using a browser or a mouse. By doing so, they were able to interact with the WBI program used in this study. Finally, 65 students from various subject areas, including business, information science, and mathematics, volunteered to participate in this study. The sample was evenly divided between male (N=32) and female (N= 33).

The empirical study consists of the following three steps:

- 1) *Identification of Cognitive Styles:* The participants' cognitive styles were classified into Holist or Serialist by Ford's Study Preference Questionnaire (SPQ), which had been used in several previous studies (e.g. Ford and Chen, 2000; 2001). The SPQ is an 18-item inventory for assessing students' learning strategies. The participants were provided with two sets of statements, one on the left and the other on the right, which they were then asked to indicate their degree of agreement with either statement, or to indicate no preference (Ford, 1985). If participants agreed with over half of the statements related to Holists, they were classified as Holists. If they agreed with over half of the statements related to Serialists, they were classified as Serialists. If they agreed with an equal number of the statements related to Holists and Serialists, they were classified as Intermediate. According to these recommendations, the sample of 65 participants was evenly distributed, with 33 Holists and 32 Serialists, but there were no intermediate students.
- 2) *Interaction with the WBI program:* The participants were asked to explore the subject content of the WBI program for approximately 90 minutes. The subject content of the WBI emphasized on the practical skills of designing web pages, in this case, *How to use HTML*. In order to examine the preferences of different cognitive styles, the WBI program provided multiple navigational tools for each participant, including an alphabetical index, a hierarchical map, a main menu, section buttons and hypertext links within the text. Thus, the participants were given the freedom to explore the instructional material in the way preferred by them. Figure 1 illustrates the design of the WBI program.
- 3) *Completion of a Questionnaire:* The participants were requested to fill out a questionnaire so that their perceptions to the use of the WBI program can be identified. This instrument was chosen because it has the potential to collect cognitive and affective data quickly and easily (Kinshuk, 1996). The questionnaire was specifically used to examine the perceptions of different cognitive style groups. Therefore, we decided to design a questionnaire specifically for this study, instead of using existing questionnaire. The reliability of the questionnaire was found to be acceptable ($\alpha=0.81$). The questionnaire consisted of 20 closed statements, which were designed to gathering specific quantitative information about students' comprehension, preferences, and

satisfaction or dissatisfaction with the WBI program, including content presentation; interaction styles; functionality and usability; and difficulties and problems. All statements used a five-point Likert Scale consisting of: ‘strongly agree’; ‘agree’; ‘neutral’; ‘disagree’; and ‘strongly disagree’. The participants were required to indicate agreement or disagreement with each statement, by placing a check mark at the response alternative that most closely reflected their opinion.

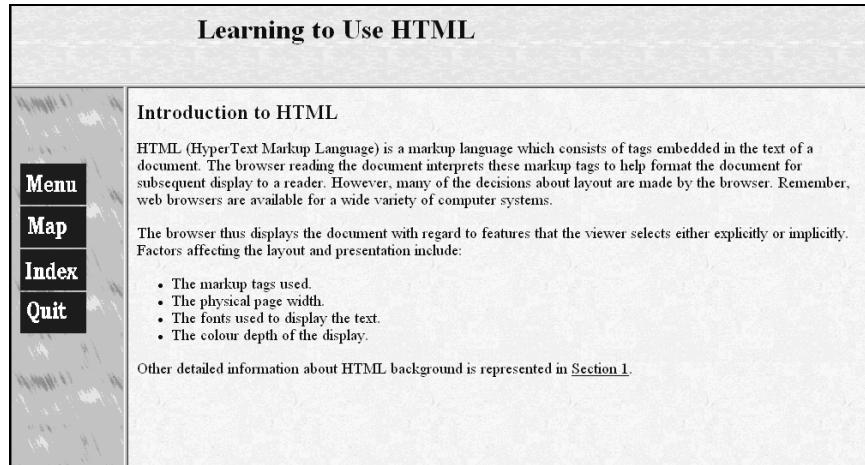


Figure 1: The WBI program

Data Analysis

This study examines how cognitive styles affect the students’ preferences for the design of the WBI program. The students’ cognitive styles were recognized as Holists and Serials with Ford’s SPQ. Their preferences were identified with a questionnaire, which included 20 statements. The responses of the statements, where appropriate, were scored as 5 for “strongly agree”, through to 1 for “strongly disagree”. As described in Introduction, data mining was applied to analyze the students’ responses to the questionnaire. The data mining process included two different stages. The first stage involved classifiers from two different families, Bayesian Networks (BN) and Nearest Neighbour (NN). These two families were chosen because of their different natures. BN use conditional probability distributions to identify the relationship between a feature and a targeted variable. On the other hand, KNN focuses on features and instances that are deemed the ‘closest’ by an imposed distance metric (Gammerman, 1997). Three of the most widely-used classifiers from each of the two classifier families were used (Table 2). These classifiers were then used to select the most relevant questions (features) from the questionnaire that were related to the Serialist/Holist cognitive style dimension.

Table 2: Two Classifier Families

| Classifier Family | Classifier |
|------------------------|---|
| Bayesian Networks (BN) | Bayesian Network (BNC) Naive Bayes (NB) Averaged-One-Dependent Estimates (AODE) |
| Nearest Neighbour (NN) | Nearest Neighbour (NNC) k-Nearest Neighbour (KNN) k-Star (K*) |

In the second stage, the resulting six feature sets selected in the first stage were then used to build decision trees. Three of the most popular decision tree algorithms, including Classification and Regression Trees (CART, Breiman et al., 1984), C4.5 (Quinlan, 1993) and CN2 (Clark and Nibbet, 1989), were used to create decisions trees. They have been chosen because they are among the most popular, the most established and the best tested in previous research (e.g., Kim, et al., 2002).

To identify a reliable decision tree, cross validation was used to assess the accuracies of the decision trees produced with the aforesaid three decision tree algorithms. This is due to the fact that cross validation is a generally applicable

and very useful technique for accuracy estimation of data mining tasks. In particular, it is widely applied for decision tree induction (Blockeel and Struyf, 2002). It consists of partitioning a data into complementary subsets, performing the analysis on one subset (called the training set), and validating the analysis on the other subset (called the testing set). Multiple rounds of cross validation are performed using different partitions and then the accuracy results from each round are then averaged to produce the estimated accuracy. In this study, ten rounds were conducted because this value has been widely used when applying cross validation to different data mining tasks (Ruiz, Riquelme and Aguilar-Ruiz, 2006). According to the results of the cross validation, the decision tree with the highest accuracy was then used to illustrate how students' cognitive style affects their preferences for the WBI.

Results and Discussion

The following section describes the results obtained from the method described in Section 3.2. Two steps were followed to obtain these results. Firstly, six classifiers from two different families were used to produce six feature sets containing only the most relevant features to Holists/Serialists. Secondly, three different decision tree algorithms were applied to produce decision trees, among which the one with the highest accuracy was then used to illustrate the effects of cognitive styles on learners' preferences for WBI.

Feature Selection

Three classifiers each from both the Bayesian Network (BN) and Nearest Neighbour (NN) families were used to obtain six feature sets. As shown in Table 3, these six classifiers selected different sets of relevant feature. The Bayesian Network classifier (BNC) selected the fewest relevant features ($n=7$), whereas the Nearest Neighbour classifier (NNC) selected the most relevant features ($n=13$). Such results contribute to the understanding of the differences among these classifiers.

Out of the twenty original features, 19 of them were selected as relevant by one or more of the classifiers. More specifically, only Q13 ("Too many options let me feel confused which options I wanted") was not selected by all of the classifiers. The feature that was commonly selected as relevant by all six classifiers was Q9 ("It is hard to use backward/forward buttons"). Three other features, Q2 ("Examples given in this tutorial are not practical"), Q14 ("After using this system I can easily use my knowledge to design home pages") and Q18 ("It is easy to find a route for a specific task with the index") were also commonly selected as relevant by five out of the six classifiers. As these four features were commonly selected by the majority of classifiers, this suggests that they are the most relevant in distinguishing the preferences of Holists and Serialists. In other words, back/forward buttons, examples and indexes are important features that need to be considered in the design of WBI to ensure that the different needs of Holists and Serialists can be accommodated.

Table 3: Relevant features selected

| No | Description | Bayesian Network Classifiers | | | Nearest Neighbour Classifiers | | |
|----|---|------------------------------|----|------|-------------------------------|-----|----|
| | | BNC | NB | AODE | NNC | KNN | K* |
| 1 | It is difficult to learn the basics of HTML using this tutorial without the help of a person. | | | | X | X | |
| 2 | Examples given in this tutorial are not practical. | X | X | X | X | | X |
| 3 | I felt the structure of this tutorial is not clear. | | | X | | | |
| 4 | I sometimes got lost because the buttons made me feel confused. | | | | | X | |
| 5 | I spent a lot of time getting to know how to use this tutorial. | | | | X | | |
| 6 | I would have found it more helpful to be given a suggested route through this tutorial. | | X | | X | X | X |
| 7 | I would like to have more examples. | | X | | X | | X |

| | | | | | | | |
|--------------|---|----------|-----------|----------|-----------|----------|-----------|
| 8 | I would prefer to learn from human tutor than from this tutorial. | | | | X | | |
| 9 | It is hard to use back/forward buttons. | X | X | X | X | X | X |
| 10 | The information provided by the map is too superficial. | X | X | X | | X | |
| 11 | The links provided in this tutorial help me discover relationships between different topics | | | X | X | | X |
| 12 | The map in this tutorial gives a meaningful framework of HTML. | X | X | | | | |
| 13 | Too many options let me feel confused which options I wanted. | | | | | | |
| 14 | After using this system I can easily use my knowledge to design home pages. | X | X | | X | X | X |
| 15 | I found it hard to select relevant information using the map. | X | X | X | | X | |
| 16 | I like the fact that it allowed me to learn topics in any order. | X | X | X | X | | |
| 17 | I like the fact that this tutorial allowed me to work at my own pace and direction. | | | | X | | X |
| 18 | It is easy to find a route for a specific task with the index. | | X | X | X | X | X |
| 19 | This tutorial can be used sufficiently well without any instructions. | | | X | | | X |
| 20 | I felt difficult to browse pages containing texts and links in the same pages. | | | | X | | X |
| Total | | 7 | 10 | 9 | 13 | 8 | 10 |

Classification

The classification stage involved two steps. Initially, three different decision tree algorithms were used to create decision trees. Secondly, the decision tree with the highest accuracy (the percentage to which how well the decision tree was able to accurately predict students' cognitive styles) was employed to illustrate how cognitive style affected students' preferences for the use of WBI programs.

Finding the most accurate feature set

Three decision tree algorithms, C4.5, CART and CN2, were used to produce decision trees with the six feature sets identified in the previous stage. In order to find the most accurate decision tree, firstly the average accuracy for each algorithm was calculated (Table 4). The idea here was that it was assumed that the decision tree with the highest average accuracy would be the best one to predict the preferences of each cognitive style. Although all three algorithms produced high accuracies, CN2 is the algorithm which produced the decision trees with the highest average accuracy.

Table 4: Classification Accuracies for the Feature Sets

| Feature Set | Classification Algorithm | | |
|-------------------------|--------------------------|--------------|-----------|
| | C4.5 | CART | CN2 |
| BN | 70.77 | 79.83 | 80 |
| NB | 70.77 | 76.79 | 80 |
| AODE | 67.69 | 87.74 | 80 |
| NN | 72.31 | 76.79 | 80 |
| KNN | 70.77 | 76.79 | 80 |
| K* | 70.77 | 76.79 | 80 |
| Average Accuracy | 70.51 | 79.12 | 80 |

After identifying the algorithm which can produce the decision trees with the highest average accuracy, the most accurate feature set was then identified from those classified with the CN2 algorithm. Somewhat surprisingly, CN2 classified all six feature sets with the same accuracy of 80%.

Illustrate student preferences with a decision tree

As the most accurate algorithm (CN2) classified all feature sets with the same accuracy (80%), it was necessary to examine the decision trees produced using these feature sets. Five different trees were produced, with the NN and KNN feature sets producing an identical tree (Figure 2), which probably reveals the underlying structure of the students' preferences revealed in this study. Moreover, this decision tree was found to include all of the information found by the other classifiers. Therefore, this decision tree is assumed to be the most accurate representation of preferences for all of those involved in the study. For this reason, this decision tree was chosen to illustrate the preferences of Holists and Serialists.

As showed in Figure 2, seven features are considered in the chosen tree to distinguish between the preferences of Holists and Serialists. In particular, three features illustrate this difference quite clearly: Q9, Q11 and Q6. Regarding Q9, learners that strongly agree or agree with "*It is hard to use back/forward buttons*" are classified as Holists. Conversely, Serialists show opposite opinions. Additionally, learners who strongly agree with Q11 ("*The links provided in this tutorial help me discover relationships between different topics*") are classified as Holists and those who strongly disagree with this are classified as Serialists. Furthermore, those users who strongly disagree with Q6 ("*I would have found it more helpful to be given a suggested route through this tutorial*") are classified as Holists whereas those that agree with this are classified as Serialists.

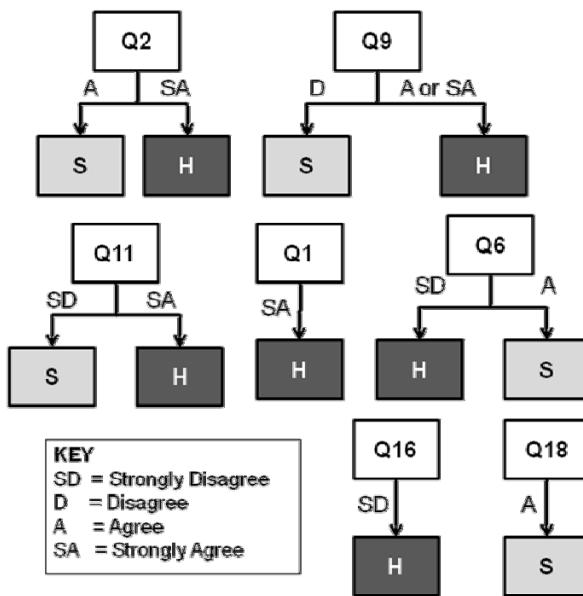


Figure 2: CN2 Decision Tree for the most accurate feature set (NN/KNN)

The results of these three features indicate that Serialists and Holists have different preferences for their navigational styles. The former prefer to follow a linear pattern by having a suggested route or looking at the subject content step-by step with back/forward buttons. Conversely, the latter tend to take a non-linear pattern by 'jumping' between different levels of subject contents with hypertext links (Jonassen and Grabowski, 1993) so they find it difficult to follow a sequential method to locate information with back/forward buttons. This is in agreement with previous literature as Serialists prefer to be given a suggested browsing pattern to direct them while Holists generally tend to use a complex linking system to find their own way (Ford et al, 1999).

The aforementioned findings demonstrate that Holists and Serialists have different preferences for the design of the WBI program. However, they share a similar preference for Q2. If a learner strongly agrees that the examples given in this tutorial are not practical, they can be a Holist or Serialist. In other words, both types of learners consider that

the examples need to be more practical. This is probably because practical examples can help them transfer knowledge into an activity (Ford & Chen, 2000). Thus, providing a concrete example may be a useful way to enhance student learning within WBI, regardless of whether they are Holists or Serialists.

Converting the Decision Tree into Decision Rules

Table 5 shows the list of rules converted from the chosen decision tree in Figure 2. As shown in this table, different rules are associated with the responses of different cognitive styles and each cognitive style is associated with more than one rule. There are six rules for Holists and five rules for Serialists. These rules can be applied to automatically suggest a student's cognitive style based on his/her preferences, which could be used to support the development of personalized WBI programs that accommodate the needs of each individual.

Table 5: Decision Rules for Holists and Serialists

| Cognitive Style | Decision Rule |
|-----------------|---|
| Holist | If learner strongly agrees that it is difficult to learn the basics of HTML without the help of a person, they are Holist. |
| | If learner strongly agrees that the examples given in this tutorial are not practical, they are Holist. |
| | If learner strongly disagrees that they would have found it more helpful to be given a suggested route through the tutorial, they are Holist. |
| | If learner strongly agrees or agrees that it is hard to use backward/forward buttons, they are Holist. |
| | If the learner strongly agrees that the links provided in this tutorial helped to discover the relationships between topics, they are Holist. |
| Serialist | If the learner strongly disagreed to the fact that the tutorial allowed them to learn topics in any order, they are Holist. |
| | If learner agrees that the examples given in this tutorial are not practical, they are Serialist. |
| | If learner agrees that they would have found it more helpful to be given a suggested route through the tutorial, they are Serialist. |
| | If learner strongly disagrees that it is hard to use backward/forward buttons, they are Serialist. |
| | If the learner strongly disagrees that the links provided in this tutorial helped to discover the relationships between topics, they are Serialist. |
| | If the learner agrees that it is easy to find a route for a specific task within the index, they are Serialist. |

Implications for System Design

The decision rules provided in Table 5 suggest that Holists and Serialists do demonstrate different preferences for the design of WBI programs. This implies that WBI programs should be developed to support the unique needs of each cognitive style group. More specifically, the WBI programs should offer multiple functionalities to accommodate the different needs of Holists and Serialists. In this section, we discuss some design solutions to support the needs of Holists and Serialists.

Design for Holists

Holists felt difficult to use this WBI program without the help of a person. In other words, there is a need to provide Holists with additional human support. To address this issue, instructors can create an email list that includes all the email addresses of the students taking the same WBI program so the students can discuss their problems and share their experience with their classmates by using this email list. Additionally, Holists thought that the links can help them discover the relationships between topics so the WBI program should present rich links to them. However, rich links may increase cognitive overhead. Thus, there is a need to use annotation, which provides additional information

about the destination of a link prior to selection (Hohl, Becker, and Gunzenhauser, 1996). Such annotation can work as a visual cue to help students decide whether the link should be followed with profit or should not yet be followed.

Design for Serialists

Unlike Holists, Serialists prefer to have a suggested route so backward /forward buttons may be a useful way for them to look for information sequentially. To help them find backward /forward buttons easily, these buttons should be clearly labeled and they are consistently located in a same place of every page. The other way to provide a suggested route for Serialists is direct guidance, which offers straightforward advice or instruction to guide students' actions. More specifically, the WBI programs offer a link or button which leads to a suggested page to read next (Brusilovsky and Millán, 2007). On the other hands, rich links may not be suitable to Serialists so we may need to use links hidden to restrict navigation space for them. More specifically, links hidden can help Serialists easily to identify which pages should be visited at the given moment and which should not so that they can be protected from the complexity of the unrestricted hyperspace.

Concluding Remarks

Understanding students' preferences is useful to both teaching and learning (Graf, Kinshuk, and Liu, 2009). The aim of this study was to investigate how cognitive style affects learners' preferences for the design of WBI programs. In using a data mining approach, the findings suggest that a learner's cognitive style tend to determine their preferences for the design of the WBI programs. Holist learners preferred the provision of additional teaching support in the form of a person. In addition, Holists also preferred to have hypertext links within the subject content that allowed them to find relationships between topics. On the other hand, Serialists preferred a suggested route through the subject content, the facilities of back/forward buttons and an index. However, both Holists and Serialists were found to appreciate practical examples.

The contributions of this study lie within its theory and methodology. In terms of the former, this study strengthens the understanding of the differences in preference between Holists and Serialists. However, it was very small-scaled though there is no minimum limit for the classifiers used in this study. Further work needs to be undertaken with a larger sample to provide additional evidence especially when a data mining approach is used to conduct data analyses. In terms of the latter, we do not only proposes a novel data mining approach that integrates feature selection and decision trees for effectively selecting and visualizing the most relevant features within a dataset, but also contributes to the knowledge of the differences among the six classifiers used for feature selection and those among the three decision tree algorithms applied for classification in this study. Nevertheless, only two families of classifiers and three decision tree algorithms were considered in this study. Such limited classifiers and decision tree algorithms may cause some bias. Thus, there is a need for further works to take into account more types and families of classifiers or other types of classification algorithms to enhance the effectiveness of this data mining approach.

Acknowledgements

This work is partially funded by National Science Council, Taiwan, ROC (NSC 98-2511-S-008-012- MY3; NSC 99-2511-S-008 -003 -MY2; NSC 99-2631-S-008-001).

References

- Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press.
- Blockeel, H., & Struyf, J. (2002). Efficient Algorithms for Decision Tree Cross-validation. *Journal of Machine Learning Research*, 3, 621-650.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth.

- Brotherton, J. A., & Abowd, G. D. (2004). Lessons Learned From eClass: Assessing Automated Capture and Access in the Classroom. *ACM Transactions on Computer-Human Interaction*, 11(2), 121-155.
- Brusilovsky, B., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. *Lecture Notes in Computer Science*, 4321, 3-53.
- Chang, L.Y., & Chen, W.C. (2005). Data Mining of Tree-Based Models to Analyze Freeway Accident Frequency. *Journal of Safety Research*, 36, 365-375.
- Chen, J., Li, Q., Wang, L., & Jia, W. (2005). Automatically generating an e-textbook on the web. *World Wide Web*, 8(4), 377-394.
- Chen, S. Y., & Liu, X. (2008). An Integrated Approach for Modeling Learning Patterns of Students in Web-based instruction: A Cognitive Style Perspective. *ACM Transactions on Computer-Human Interaction*, 15(1), Article 1.
- Chen, S. Y., & Macredie, R. D. (2004). Cognitive Modelling of Student Learning in Web-based Instructional Programmes. *International Journal of Human-Computer Interaction*, 17(3), 375-402.
- Chen, Y. L., Hsu, C. L., & Chou, S. C. (2003). Constructing a multi-valued and multi-labeled decision tree. *Expert Systems with Applications*, 25(2), 199-209.
- Clark, P., & Niblett, T. (1989). The CN2 Induction Algorithm. *Machine Learning*, 3(1), 261-283.
- Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology*, 16, 65-79.
- Ford, N. (2000). Cognitive Styles and Virtual Environments. *Journal of the American Society for Information Science*, 51(6), 543-557.
- Ford, N. (2001). The increasing relevance of Pask's work to modern information seeking and use. *Kybernetes*, 30(5/6), 603-629.
- Ford, N., & Chen, S. Y. (2001). Matching/Mismatching Revisited: An Empirical Study of Learning and Teaching Styles. *British Journal of Educational Technology*, 32(1), 5-22.
- Ford, N., & Chen, S. Y. (2000). Individual Differences, Hypermedia Navigation and Learning: An Empirical Study. *Journal of Educational Multimedia and Hypermedia*, 9(4), 281-312.
- Ford, N., & Ford, R. (1993). Towards a cognitive theory of information accessing: an empirical study. *Information Processing and Management*, 29(5), 569-585.
- Ford, N., Wilson, T.D., Foster, A. E., & Ellis, D. (1999). Information Seeking and Mediated Searching: Part 4 - Cognitive Styles in Information Seeking. *Journal of the American Society for Information Science and Technology*, 53(9), 728-35.
- Gammerman, A. J. (1997). *Machine learning: progress and prospects*. London: University of London.
- Graf, S., Kinshuk, & Liu, T.-C. (2009). Supporting Teachers in Identifying Students' Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach. *Educational Technology & Society*, 12(4), 3-14.
- Graff, M (2003). Learning from Web-based Instructional Systems and Cognitive Style, *British Journal of Educational Technology*, 34, 407-418.
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann.
- Hand, D.J., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. MIT Press.
- Harumoto, K., Nakano, T., Fukumura, S., Shimojo, S., & Nishio, S. (2005). Effective Web browsing through content delivery adaptation. *ACM Transactions on Internet Technology*, 5(4), 571-600.
- Hohl, H., Böcker, H. & Gunzenhäuser, R. (1996). Hypadapter: An adaptive hypertext system for exploratory learning and programming. *User Modeling and user adapted Interaction*, 6, 131-156.
- Huan, L. & Lei, Y. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17(4), 491-502.
- Jonassen, D.H. & Grabowski, B.L. (1993). *Handbook of Individual Differences, Learning and Instruction*. New York: Lawrence Erlbaum.
- Kim, J.K., Cho, Y.H., Kim, W.J., Kim, J.R., & Suh, J.H. (2002). A personalized recommendation procedure for Internet shopping support. *Electronic Commerce Research and Applications*, 1(1), 301-313.
- Kinshuk (1996). *Effectiveness of Intelligent Tutoring Tools Interfaces in relation to Student, Learning Topic and Curriculum Characteristics*. PhD Thesis. De Montfort University, United Kingdom.
- Kirkwood, A. (2008). Getting it from the Web: why and how online resources are used by independent undergraduate learners. *Journal of Computer Assisted Learning*, 24(5), 372-382.

- Lee, W., Chen, S.Y., Crysostomou, K., & Liu, X. (2009). Mining students' behavior in web-based learning programs. *Expert Systems with Applications*, 36(2), 3459-3464.
- Liu, H., & Kešelj, V. (2007). Combined mining of Web server logs and web contents for classifying user navigation patterns and predicting users future requests. *Data & Knowledge Engineering*, 61(2), 304-330.
- Ma, Z., Pant, G., & Sheng, O. R. (2007). Interest-based personalized search. *ACM Transactions on Information Systems*, 25(1), Article 5.
- Messick S. (1976). *Individuality in learning*. San Francisco: Jossey-Bass.
- Mitchell, T. J. F., Chen, S. Y., & Macredie, R. D. (2005). Hypermedia Learning and Prior Knowledge: Domain Expertise vs. System Expertise. *Journal of Computer Assisted Learning*, 21, 53-64.
- Pask, G. (1979). Styles and strategies of learning. *British Journal of Educational Psychology*, 46, 128-148.
- Quinlan, J.R. (1993). *C4.5: Programs for Machine Learning*. Massachusetts: Morgan Kaufmann.
- Riding, R. & Rayner, S. G. (1998). *Cognitive Styles and Learning Strategies*. London: David Fulton.
- Ruiz, R., Riquelme, J.C., & Aguilar-Ruiz, J.S. (2006). Incremental wrapper-based gene selection from microarray data for cancer classification. *Pattern Recognition*, 39, 2383-2392.
- Romero, C., & Ventura, S. (2006). *Data Mining in e-learning*. Boston: Wit Press.
- Romero, C., & Ventura, S. (2007) Educational Data Mining: A Survey from 1995-2005. *Expert Systems with Applications*, 33(1), 135-145.
- Sears, A., & Jacko, J. A. (2009). *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. New York: Taylor & Francis.
- Tang, C., Yin, H., Li, T., Lau, R., Li, Q., & Kilis, D. (2000). Personalized courseware construction based on web data mining. *Paper presented at the first international conference on Web Information Systems Engineering*, June 19-20, Hong Kong, China.
- Urtubia, A., Perez-Correa, J.R., Soto, A., & Pszczolkowski, P. (2007). Using data mining techniques to predict industrial wine problem fermentations. *Food Control*, 18(1), 1512-1517.
- Witkin, H.A. (1976). Cognitive style in academic performance and in teacher-student relations. In S. Messick (Ed.), *Individuality in learning: Implications of cognitive style and creativity for human development*. San Francisco: Jossey-Bass.
- Yang, J., & Olafsson, S. (2006). Optimization-based feature selection with adaptive instance sampling. *Computers and Operations Research*, 33(11), 3088-3106.

Teaching Online: Tools and Techniques, Options and Opportunities

(Book Review)

Reviewer:

David Leal Cobos

Teacher of English at Secondary Education
Colegio los Ángeles (Getafe)
Calle Dehesillas 8, Portal 4, 4ºD
Fuenlabrada, Madrid, Spain
ancolla@hotmail.com

Textbook Details:

Teaching Online: Tools and Techniques, Options and Opportunities
Nicky Hockly and Lindsay Clandfield
Delta Publishing, pp. 112
ISBN 978-1-905085-35-4

Teaching Online. Tools and Techniques, Options and Opportunities belongs to the praised Delta Teacher Development Series which has become one of the most innovative and important resources for ELT professionals in the past few years. This book is designed to reach a wide variety of readers immersed in the teaching of the English language at all levels. Not only will it be useful for professionals who are already experts in online tools and environments, but also for educators whose teaching methodologies have not yet included online or computer based activities and tasks. Thus, this work is aimed at a wide range of audience, from beginners who may not feel sure about the use of computer technologies in class, to advanced users who look forward to learning new teaching practices related to this field. *Teaching Online. Tools and Techniques, Options and Opportunities* is also commendable because it is clear, reassuring and motivating because it removes the terms “complexity” and “difficulty” from the general understanding of online teaching.

The book is divided into three main sections (A, B, and C) that cover the basic rudiments to start an online course, a lot of practical activities to be carried out by both teachers and students, and possibilities for further online development. Section A deals with the growing importance of online teaching and learning, as well as the different possibilities to establish courses depending on the level of online work. The reader is informed of the differences between online learning and blended learning, (combining both face-to-face and online learning elements) and its advantages and disadvantages. The authors then highlight the importance of organisation regarding different aspects such as the number of classes a week, duration of the course, and the type of students. Having taken into account these elements, the reader is guided to reflect on the amount and purpose of online components in the course, the role of the teachers themselves, different learning scenarios and options to establish the most suitable type of course depending on the variables mentioned beforehand (mainly face-to-face, half-and-half, and mainly online).

Section A also highlights the importance of the right choice of software for the online sections. The use of a Virtual Learning Environment (VLE) is mainly recommended to deliver the online course content, apart from other course site tools such as Social Networking Sites, Wikis and Discussion Groups. These “Web 2.0” tools are expanded and studied in depth where the reader is presented with a wide array of free activity tools ranging from the traditional blogs and word clouds to the more advanced screen capture tools and shared whiteboards. Finally, this section informs the reader of the importance of the human factor when working on the online part of a course. Aspects of meeting and greeting, setting of objectives, celebration of achievements, deadlines, manners and netiquette are brought to the front so that the potential online teacher knows how to act and establish the basic grounds for online co-existence.

Section B offers a varied set of practical activities with possible variations, follow-up tasks, and feedback options designed to cover the whole duration of the course. To attain the goal of clarity and easy accessibility, the authors have divided this section into five chapters. The first subsection, “The Starting Line”, is designed to facilitate the process of the teacher getting to know the learners online and the learners getting to know themselves. In addition, the reader is given effective advice on how to make this initial period of the course easier by means of, for example,

creating a strong sense of community. Then, the authors propose twelve different activities with clear outlines and purposes which makes use of the different tools already explained in Section A. The second chapter, “Reading and Writing Online”, offers twenty-one activities for these two skills, which are usually thought to be the easiest to teach online because they are supposed to be individual processes. The authors do not agree with this point of view and, thus, make an emphasis on the most difficult aspect of teaching Reading and Writing online: the engagement of the students in interactive, encouraging tasks oriented to both skills and the reinforcement of the learners’ community by providing private and public places of contact.

The third chapter, “Listening and Speaking Online”, is arguably the most successful section in the book. Teaching the skill of Listening online does not pose any serious difficulty, since there are lots of available resources and materials for teachers and learners. However, the skill of Speaking is really challenging because the most typical set of activities are “listen and repeat” ones. The authors propose eighteen activities (from less demanding to more demanding ones) to cover both skills starting with two listening tasks which use generic sites. Afterwards, the book focuses on individual speaking activities and, finally, group speaking tasks. The fourth chapter, “Language and Evaluation Online”, suggests language work and assessment methods that can be performed online individually or as a group in meaningful and communicative ways. Nineteen activities are proposed for such objectives. They are easily adaptable and implementable to any course regardless of its level and content. The fifth chapter, “The Finishing Line”, offers four activities aimed to provide teachers with tasks to draw their course to an end with a sense of closure. In fact, the final activity is designed to make learners reflect on their learning process and proposes teachers to encourage them by suggesting online resources to continue learning.

Finally, Section C remarks that the improvement of the internet has provided teachers with huge development opportunities worldwide. As the authors state, knowledge is not only in the hands of a few experts and we can share tips, strategies, techniques, etc. with colleagues working in different countries with different contexts. As a means to facilitate this educational exchange, the book lists six different discussion groups and recommends teachers to join and contribute to them. The authors, then, investigate the field of Development Courses and give readers some tips to select the most appropriate one. Furthermore, online conferences are still a key means of teachers’ development for the authors and they include their favourite ones, ranging from a one-hour webinar to a four-day international conference. Blogging and Micro-Blogging are also mentioned in this section along with their differences and the authors’ favourites. E-Portfolios are explained in-depth regarding their intended audience, the tool the teacher is going to use, its sharing, and distribution. Lastly, Personal Learning Networks (PLN) are also part of this final section and are dealt with in a clear accessible way, so that readers do not get lost in this sea of technology and development.

In conclusion, *Teaching Online. Tools and Techniques, Options and Opportunities* is highly recommendable to any person dealing with the teaching of English. Its clear structure and its motivating tone engages readers who will be willing to incorporate these tools, techniques and activities in their teaching practice, since they encompass the four skills and even go beyond the mere teacher-learner relationship. Moreover, the recommended activities are highly adaptable to any level, and they facilitate the creation of new activities and class dynamics by teachers themselves. It is remarkable the ease with which the authors tackle this topic of online teaching and the demystification it is subjected to in the book, making it accessible to anyone curious enough to explore a new field that is not as complex as it appears to be. Finally, although oriented to English language teaching, it cannot be denied that it is also applicable to the teaching of any language, since the activities are not only restricted to one language. To sum up, this book is a must-read for anyone interested in incorporating computer technology and the internet in class.

<http://www.ifets.info>

ISSN: 1436-4522 (online)
ISSN: 1176-3647 (print)