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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.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:

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

Editorial Assistant
Barbara Adamski, Athabasca University, Canada.

Associate editors
Nian-Shing Chen, National Sun Yat-sen University, Taiwan; Alexandra I. Cristea, Technical University Eindhoven, The Netherlands; John Eklund, Access Australia Co-operative Multimedia Centre, Australia; Vladimir A Fomichev, K. E. Tsiolkovsky Russian State Tech Univ, Russia; Olga S Fomicheva, Studio "Culture, Ecology, and Foreign Languages", Russia; Piet Koomers, 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; Luis Anido-Rifon, University of Vigo, Spain; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bullen, University of British Columbia, Canada; Tak-Wai Chan, National Central University, Taiwan; Sherry Chen, Brunel University, United Kingdom; Darina Dicheva, Winston-Salem State University, USA; Robert Farrell, IBM Research, USA; Brian Garner, Deakin University, Australia; Roger Hartley, Leeds University, UK; J R Isaac, National Institute of Information Technology, India; Mohamed Jenni, University of Tunis, Tunisia; Paul Kirschner, Open University of the Netherlands, The Netherlands; William Klemm, Texas A&M University, USA; Rob Koper, Open University of the Netherlands, The Netherlands; Rudy Lelouche, Université Laval, Canada; David McConnell, Lancaster University, UK; Rory McCread, Athabasca University, Canada; David Merrill, Brigham Young University - Hawaii, USA; Marcello Milrad, Växjö University, Sweden; Riichiro Mizoguchi, Osaka University, Japan; Permanand Mohan, The University of the West Indies, Trinidad and Tobago; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, The University of Electro-Communications, Japan; Thomas C. Reeves, The University of Georgia, USA; Gilly Salmon, University of Leicester, United Kingdom; Norbert M. Seel, Albert-Ludwigs-University of Freiburg, Germany; Timothy K. Shih, Tamkang University, Taiwan; Yoshiaki Shinjo, Nippon Institute of Technology, Japan; Kevin Singley, IBM Research, USA; Brian K. Smith, Pennsylvania State University, USA; J. Michael Spector, Florida State University, USA; Chin-Chung Tsai, National Taiwan University of Science and Technology, Taiwan; Stephen J.H. Yang, National Central University, Taiwan.

Assistant Editors
Sheng-Wen Hsieh, Far East University, Taiwan; Taiyu Lin, Massey University, New Zealand; Kathleen Luchini, University of Michigan, USA; Dorota Mularczyk, Independent Researcher & Web designer; Carmen Padrón Nápoles, Universidad Carlos III de Madrid, Spain; Ali Fawaz Shareef, Massey University, New Zealand; Jarkko Suonen, University of Joensuu, Finland.

Executive peer-reviewers
http://www.ifets.info/

Subscription Prices and Ordering Information
For subscription information, please contact the editors at kinshuk@ieee.org.

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


Guidelines for authors

Submissions are invited in the following categories:

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

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

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

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

Submission procedure

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

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

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

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

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

Special issue articles

Knowledge infrastructure of the future (Guest Editorial)
   Nian-Shing Chen and Chin-Chung Tsai
   1-4

National Program for e-Learning in Taiwan
   Maiga Chang, Chin-Yeh Wang and Gwo-Dong Chen
   5-17

An Evaluation of the ELNP e-Learning Quality Assurance Program: Perspectives of Gap Analysis and Innovation Diffusion
   Ming-Puu Chen
   18-33

The Model of Strategic e-Learning: Understanding and Evaluating Student e-Learning from Metacognitive Perspectives
   Meng-Jung Tsai
   34-48

An Automatic Multimedia Content Summarization System for Video Recommendation
   Jie Chi Yang, Yi-Ting Huang, Chi Cheng Tsai, Ching I Chung and Yu Chieh Wu
   49-61

A Study on the Learning Efficiency of Multimedia-Presented, Computer-Based Science Information
   Ying-Hua Guan
   62-72

Designing a Syntax-Based Retrieval System for Supporting Language Learning
   Nai-Lung Tsao, Chin-Hwa Kuo, David Wible and Tsung-Fu Hung
   73-81

A Knowledge-based Approach to Retrieving Teaching Materials for Context-aware Learning
   Wen-Chung Shih and Shiian-Shyong Tseng
   82-106

Ubiquitous Performance-support System as Mindtool: A Case Study of Instructional Decision Making and Learning Assistant
   Hsinyi Peng, Po-Ya Chuang, Gwo-Jen Hwang, Hui-Chun Chu, Ting-Ting Wu and Shu-Xian Huang
   107-120

Content-Free Computer Supports for Self-Explaining: Modifiable Typing Interface and Prompting
   Chih-Yueh Chou and Hung-Ta Liang
   121-133

Examining the Factors Influencing Participants’ Knowledge Sharing Behavior in Virtual Learning Communities
   Irene Y.L. Chen, Nian-Shing Chen and Kinshuk
   134-148

Applications of a Time Sequence Mechanism In the simulation Cases of a web-based Medical Problem-Based Learning System
   Lih-Shyang Chen, Yuh-Ming Cheng, Sheng-Feng Weng, Yong-Guo Chen and Chyi-Her Lin
   149-161

The information commitments toward web information among medical students in Taiwan
   Jyh-Chong Liang and Chin-Chung Tsai
   162-172

Full length articles

Automatic Generation of Just-in-time Online Assessments from Software Design Models
   Imran A. Zualkernan, Salim Abou El-Naaf, Maria Papadopoulos, Budoor K. Al-Amoudi and Charles E. Matthews
   173-192

Main Barriers and Possible Enablers of ICTs Integration into Pre-service Teacher Education Programs
   Yusel Goktas, Soner Yildirim and Zahide Yildirim
   193-204

Utilizing Computer-mediated Communication Tools for Problem-based Learning
   Hao-Chang Lo
   205-213
<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examining Applicants for Admission and Completion of an Online Teacher Certification Program</td>
<td>214-229</td>
</tr>
<tr>
<td>Jon J. Denton, Trina J. Davis, Robert M. Capraro, Ben L. Smith, Lynn Beason, B. Diane Graham and R. Arlen Strader</td>
<td></td>
</tr>
<tr>
<td>A Distributed Online Curriculum and Courseware Development Model</td>
<td>230-248</td>
</tr>
<tr>
<td>Pinar Onay Durdu, Nese Yalabik and Kursat Cagiltay</td>
<td></td>
</tr>
<tr>
<td>Asynchronous discussion forums: success factors, outcomes, assessments, and limitations</td>
<td>249-257</td>
</tr>
<tr>
<td>Martin A. Andresen</td>
<td></td>
</tr>
<tr>
<td>Using webquests for oral communication in English as a foreign language for Tourism Studies</td>
<td>258-270</td>
</tr>
<tr>
<td>Jesús Garcia Laborda</td>
<td></td>
</tr>
<tr>
<td>The effects of web-based instruction navigation modes on undergraduates’ learning outcomes</td>
<td>271-284</td>
</tr>
<tr>
<td>Yu-Chang Hsu, Huifen Lin, Yu-Hui Ching and Francis M. Dwyer</td>
<td></td>
</tr>
<tr>
<td>Information society needs of managers in a large governmental organisation</td>
<td>285-297</td>
</tr>
<tr>
<td>Elizabeth Broos and Johannes C Cronjé</td>
<td></td>
</tr>
</tbody>
</table>
Knowledge infrastructure of the future (Guest Editorial)

Nian-Shing Chen
National Sun Yat-sen University, Taiwan // nschen@faulty.nsysu.edu.tw

Chin-Chung Tsai
National Taiwan University of Science & Technology, Taiwan // cct sai@mail.ntust.edu.tw

This special issue of *Educational Technology & Society* aims at giving the reader a highlight of current e-Learning research in Taiwan. The articles presented in this issue features the best research papers from the Technology Enhanced Learning (TELearn2007) conference held at the National Central University, Taiwan, in July 2007.

In general, to be able to offer an online course, one must first to have a learning platform or so called Learning Management System (LMS) to use, and then upload prepared learning contents for the course onto the platform, at last conduct learning activities by using the functions provided by the learning platform. However, there are two more unique features that online can do much better than its physical counterpart which are learning community and knowledge management.

Due to no time and space barriers and flexible of carry on asynchronous and synchronous learning activities, collaborative learning, peer learning and active social learning can be easily realized by running a successful learning community. The rapid developments of Web 2.0 applications and the popularity of social software further validated this argument. The challenge is how to run a successful learning community in an online learning environment; most teachers are still lacking skills and experiences and many issues are yet to be explored.

The two major drawbacks of physical teaching & learning are difficult to document teaching & learning portfolios and not easy to reuse and value-added previous materials. An online course is delivered in the form of digital format, every piece of digital material can be archived and every activity can be tracked and logged. Therefore, it is very important to incorporate knowledge management in running an online course, such that elegant materials created by previous students can be accumulated from semester to semester for beneficial later students. Teachers can document complete digital teaching portfolios for better reuse or even value-added of their online courses to enhance teaching efficiency and performance.

There are three main roles in any educational settings, which are instructor, learner and administrator; this is the same as for e-learning. Let’s propose a dual modals framework of e-Learning for better representing the context of e-Learning as shown in Figure 1. The first modal contains learning platform, learning content, learning activity, learning community and knowledge management which lying on the outer circle; the second modal contains instructor, learner and administrator which lying on the inner circle. To achieve a high quality e-learning education, all the people playing these three roles need to develop professional knowledge and know-how of how to best fulfill their roles by interacting with the five critical components surrounding around them in the context. We will use this framework to categorize the papers in this issue.

The evolution of the Internet applications has been gone through cyber community, e-learning and knowledge management. We are now in the period of transforming from knowledge management to innovation management. All these four types of applications, that is, community management, learning management, knowledge management and innovation management, need well-designed Internet application systems to carry. By examining from the system perspective, community management system and learning management system are already very mature; knowledge management system is moderate mature; innovation management system is far beyond mature yet. It seems that to develop a mature innovation management system still needs a lot of research and effort to be devoted in the future.

It is true but we observed that one cannot do good e-learning without doing learning community; one cannot do good knowledge management without doing e-learning. The same principle works for innovation management, which means one cannot do good innovation management without doing knowledge management. We can infer from these observations that a well-designed learning management system should include functions supported by the
community management system; and a well-designed knowledge management system should include functions supported by the learning management system; and a well-designed innovation management system should include functions supported by the knowledge management system. In summary, \( F(\text{CMS}) \subset F(\text{LMS}) \subset F(\text{KMS}) \subset F(\text{IMS}) \).

*Figure 1. A dual modal framework of e-Learning*

*Figure 2. The evolution of Internet applications from time and function perspectives*
We then argue that it is not a good idea to develop an innovation management system from a scratch, there is a more efficient way to develop an innovation management system by just adding additional functions on an existing knowledge management system. This concept can also be applied to do research related to community management, learning management, knowledge management and innovation management. As depicted in Figure 2, since innovation management system actually includes the functionalities supported by knowledge management system, and knowledge management system actually includes the functionalities supported by learning management system, and the same as for learning management system which actually includes the functionalities supported by community management system. That is to say, researchers can apply methodologies and methods which have been proved to be useful in solving community management and learning management related problems to explore knowledge management and innovation management issues.

Maiga Chang, Chin-Yeh Wang and Gwo-Dong Chen in their paper entitled “National Program for e-Learning in Taiwan” introduce an e-learning national program initiative in Taiwan, and the program is called the National Science and Technology Program for e-Learning. The research topics and issues studied in this national program cover almost every component described in Figure 1.

Ming-Puu Chen in the paper entitled “An Evaluation of the ELNP e-Learning Quality Assurance Program: Perspectives of Gap Analysis and Innovation Diffusion” examines the appropriateness of a nationwide quality assurance framework for e-learning from participants’ perspectives. The issues of this research almost cover every component described in Figure 1.

Meng-Jung Tsai in the paper entitled “The Model of Strategic e-Learning: Understanding and Evaluating Student e-Learning from Metacognitive Perspectives” constructs a model and develops an instrument called OLSS to present several future research directions. This is to address almost every component described in Figure 1.

Jie Chi Yang, Yi Ting Huang, Chi Cheng Tsai, Ching I Chung and Yu Chieh Wu in their paper entitled “An Automatic Multimedia Content Summarization System for Video Recommendation” explore how to effectively use videos for learning. This is to address the issues of learning content and learner as shown in Figure 1.

Ying-Hua Guan in the paper entitled “A Study on the Learning Efficiency of Multimedia-Presented, Computer-Based Science Information” explores the effects of different sorts of multimedia information on students’ science learning. This is to address the issues of learning content, instructor and learner as shown in Figure 1.

Nai-Lung Tsao, Chin-Hwa Kuo, David Wible and Tsung-Fu Hung in their paper entitled “Designing a Syntax-Based Retrieval System for Supporting Language Learning” propose a syntax-based text retrieval system for on-line language learning. This is to address the issues of learning platform, learning activity and learner as shown in Figure 1.

Wen-Chung Shih and Shian-Shyong Tseng in their paper entitled “A Knowledge-based Approach to Retrieving Teaching Materials for Context-aware Learning” propose a knowledge-based system to solve content retrieval problems of the context-aware learning. This is to address the issues of learning platform, learning content and knowledge management described in Figure 1.

Hsinyi Peng, Po-Ya Chuang, Gwo-Jen Hwang, Hui-Chun Chu, Ting-Ting Wu and Shu-Xian Huang in their paper entitled “Ubiquitous Performance-support System as Mindtool: A Case Study of Instructional Decision Making and Learning Assistant” propose a system called UPSS that can facilitate the seamless use of powerful new technologies in the school setting. It is a useful reference for those who are interested in conducting studies applying context-aware ubiquitous computing to educational contexts. This is to address almost every component described in Figure 1.

Chih-Yueh Chou and Hung-Ta Liang in their paper entitled “Content-Free Computer Supports for Self-Explaining: Modifiable Typing Interface and Prompting” investigate the effects of self-explaining on learning. This is to address the issues of learning activity, instructor and learner as shown in Figure 1.

Irene Y.L. Chen, Nian-Shing Chen and Kinshuk, in their paper entitled “Examining the Factors Influencing Participants’ Knowledge Sharing Behavior in Virtual Learning Communities” attempt to examine the factors
influencing knowledge sharing from the perspective of human behavior. This is to address the issues of learning community, knowledge management and administrator as shown in Figure 1.

Lih-Shyang Chen, Yuh-Ming Cheng, Sheng-Feng Weng, Yong-Guo Chen and Chyi-Her Lin in their paper entitled “Applications of a Time Sequence Mechanism in the Simulation Cases of a web-based Medical PBL System” analyze the organization of a computerized PBL teaching case and consider how a clinical teaching case can best be presented to the users. This is to address the issues of learning activity, instructor and learner as shown in Figure 1.

Jyh-Chong Liang, and Chin-Chung Tsai in their paper entitled “The information commitments toward web information among medical students in Taiwan” use the Information Commitment Survey (ICS) for an investigation of medical students’ standards of judging online information and their search strategies on the Web. This is to address the issues of learning activity, instructor and learner as shown in Figure 1.
National Program for e-Learning in Taiwan

Maiga Chang¹, Chin-Yeh Wang², Gwo-Dong Chen³
¹School of Computing & Information Systems, Athabasca University, Canada
²National Science and Technology Program for e-Learning in Taiwan, Taiwan
³Dept. of Computer Science and Information Engineering, National Central University, Taiwan
maiga@ms2.hinet.net, chinyea@db.csie.ncu.edu.tw, chen@csie.ncu.edu.tw

ABSTRACT
Taiwan government has initiated a five-year program since 2002: the National Science and Technology Program for e-Learning. The national program started from 2003 and was completed at the end of year 2007, involving thirteen government agencies. This paper describes the results that the national program has accomplished at its first phase, 2003 to 2007. The results include how the national program has helped enhance competitiveness of commerce; improved public welfare in Taiwan; and how the national program has stimulated the research outputs in both industry and academia.

Keywords
Taiwan National e-Learning Program, Digital learning, e-Learning development strategies, Academic and industry developments

Introduction

Many developed and developing countries have currently allocated funds and resources to encourage researches in e-learning with a view to gaining profits it can bring (Hwang, 2003). In Europe, the Commission of the European Communities (2001) announced the guideline of e-learning policy, The e-Learning Action Plan—Designing tomorrow’s education; the Secretary of Commerce in the United States (2002) published "2020 Visions—Transforming Education and Training through Advanced Technologies" report; and, Taiwan government has also planned a five-year national program for e-learning in 2002 National Science & Technology Program for e-Learning, (2002).

The national program, National Science and Technology Program for e-Learning, started from 2003. The major objective of the national program is to increase happiness among common citizens via e-learning, for example, students may feel happy if they get good grades or get permission from their favorite universities; parents may feel happy if their children have good performances at school; business managers may feel happy if they cut down the cost of human resource training; workers may feel happy if they get promotion and raise of salary; moreover, learners may feel happy if they can study the topics they’re interested in. There are still two goals for the program to accomplish: broaden knowledge among people and enhance national competitiveness.

Thirty government agencies have joined the national program, each in charge of different aspects of promoting e-learning, for example, e-learning can be applied to business and enhance competitiveness; improve public welfare; help develop new learning technologies, methodologies, systems, and tools. Regarding enhancing commercial competitiveness, the national program assigns Industrial Development Bureau of Ministry of Economic Affairs (MOEA) and Industrial Technology Department of MOEA to the conduction; improving public welfare, the national program has Council of Labor Affairs, Ministry of Education, Council for Culture Affairs, Council for Indigenous Peoples (joined since 2005), Council for Hakka Affairs (joined since 2005), Bureau of Health, and National Palace Museum (is considered one of world's great museum; Japan Aerospace Exploration Agency, 2005) perform the conduction; doing e-learning related research in both industry and academia, Industrial Technology Department of MOEA and National Science Council are involved.

Currently, there is no national e-learning program promoted by government around the world. Most large e-learning programs are supported by research organizations, non-profit organizations, and corporations; furthermore, it is rare that an e-learning program involves so many government agencies as the national program for e-learning in Taiwan does. This report reveals the way of promoting e-learning from the viewpoint of the government. Moreover, it reveals how such kind of national program succeeds in helping corporations enhance competitiveness, improving public welfare in Taiwan, and stimulating the research outputs in both industry and academia. However, there are
still some issues for the government to solve when promoting e-learning; different participating government agencies may have different viewpoints and expectations to e-learning, and the government needs to sort out the best way for agencies to work together. This paper will be a good reference and source for other countries when trying to promote e-learning.

This paper first describes the details of the national program for e-learning in Taiwan in Section 2. Section 3 shows how the five-year national program improves public welfare. Furthermore, the program has enhanced competitiveness of not only the industry in e-learning but also many other industries. Section 4 gives the examples of how the national program has enhanced economical competitiveness in different industries. Section 5 illustrates the research results in two ways: the industrial patents and the academic papers. Finally, a simple conclusion is made in Section 5.

**National Science and Technology Program for e-Learning**

Via e-learning, services, information, knowledge and policies are delivered from government agencies to the public more efficiently. Applying e-learning to the organization, corporations may cut down time and budgets in training employees, increase employees' learning efficiency, and reduce the divide in professionalism. Educational and training organizations can help learners improve learning efficiency and teachers teach more efficiently via information communication technologies (ICT) and digital technologies of e-learning. According to these benefits and technological trends, Taiwan government initiated the national science and technology program for e-learning in order to equalize the opportunity for learning among common citizens, promote e-learning industry, and increase commercial competence in Taiwan.

National Science and Technology Program for e-Learning is a five-year program since Jan. 2003. The first phase lasts from 2003 to 2007 and the National Science Council has also planned the second phase, the promotion phase (2008-2012). There are seven operational programs in ELNP, including "e-learning for everyone", "narrowing the digital divide", "mobile learning devices", "internet-based industrial park for e-learning (e-learning park)", "advanced e-learning technology R&D", "fundamental research on learning and cognition in e-learning", and "policy guidance and manpower cultivation".

Table 1 lists the total budgets for the national program from 2003 to 2007. Every agency has put all efforts to make the final budgets meet what's planned.

<table>
<thead>
<tr>
<th>Year</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned</td>
<td>25.7</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>133.7</td>
</tr>
<tr>
<td>Actual</td>
<td>20.58</td>
<td>25.61</td>
<td>24.59</td>
<td>19.63</td>
<td>20.33</td>
<td>110.74</td>
</tr>
</tbody>
</table>

*final government budgets

Thirteen government agencies were involved in the national program, the twelve government agencies were National Science Council, Industrial Development Bureau of MOEA, Industrial Technology Dept of MOEA, Ministry of Education, Ministry of National Defense (from 2003 to 2005), Bureau of Health, Council of Labor Affairs, Council for Culture Affairs, Council of Indigenous Peoples (since 2005), Council for Hakka Affairs (since 2005), Tainan County Government (from 2003 to 2005), Overseas Compatriot Affairs Commission (2007), and National Palace Museum. The seven operational programs aim at three goals: to equalize learning opportunity for people; to increase commercial competence in Taiwan; to enhance quality of academic researches and development of industrial technology in e-learning.

The government sets three goals for the national program. And many government agencies have participated in the national program, each having different subjects and emphasis. For example, Council of Labor Affairs aims at workers and the unemployed to help them get jobs; MOEA focuses on helping commercial competence through e-learning and e-learning industry get orders from corporations around the world; National Science Council emphasizes on enhancing the quality of academic researches and how to put the results in practical use. For better...
cooperation among the agencies, the national program divided the agencies into different goal groups, and each is asked to integrate expected yearly outcome and reach mutual support.

The national program divided the participating government agencies into three goal groups:

1. **equalizing opportunities for learning**
   - Ministry of Education (MOE)
   - Bureau of Health (BOH)
   - Council of Labor Affairs (CLA)
   - Council for Culture Affairs (CCA)
   - Council of Indigenous Peoples (CIP, since 2005)
   - Council for Hakka Affairs (CHA, since 2005)
   - Industrial Development Bureau of MOEA (IDB of MOEA)
   - Taichung County Government (TCG, from 2003 to 2005)
   - Overseas Compatriot Affairs Commission (OCAC, 2007)
   - National Palace Museum (NPM)

2. **enhancing commercial competence in Taiwan**
   - Industrial Development Bureau of MOEA (IDB of MOEA)
   - Council of Labor Affairs (CLA)

3. **enhancing the quality of academic researches and development of industrial technologies in e-learning**
   - National Science Council (NSC)
   - Industrial Development Bureau of MOEA (IDB of MOEA)
   - Industrial Technology Dept of MOEA (ITD of MOEA)
   - Ministry of Education (MOE)

The national program required complementary collaboration among agencies, for example, MOE developed the standard operating procedure (SOP) of instruction design for e-learning courses and materials, and this should be the reference for other agencies; CLA, BOH, and NPM, developed numerous high quality e-learning courses and materials, other agencies such as TCG could resort to it directly; CHA set up a certificate for Hakka language competence and BOH a certificate for medical personnel, CLA could apply regulations to it; IDB of MOEA built an e-learning park gathering many corporations involved in the industry and provided high-quality e-learning products and services, all agencies could resort to the products and services provided by the e-learning park.

Two of the seven operational programs, "e-learning for everyone" and "narrowing the digital divide", are relevant to the public welfare improvement; another two operational programs, "internet-based industrial park for e-learning (e-learning park)" and "advanced e-learning technology R&D", are relevant to increase business competitiveness. Furthermore, the "advanced e-learning technology R&D" and fundamental research on learning and cognition in e-learning" operational programs are relevant to the researches corresponding to industry and academia. This paper mainly focuses on the outcomes of these five operational programs.

Regarding public welfare improvement, the national program plans on six steps of execution: improving laborers' professionalism; narrowing the digital divide among laborers such as disabled workers and the unemployed; cultivating e-learning related professions; broadening knowledge among citizens, as government agencies provide chronic diseases e-learning courses and Chinese culture/artifacts e-learning courses; increasing professionalism community reconstruction designers; and increasing aboriginal people's learning opportunity. The national program provides workers continuous learning to enhance their working skills, get promotions, and increase their wage; prepares the unemployed for satisfying jobs; aids students in academic achievement and permission into the schools they want; equalizes the opportunities for learning and working among the disabled, aborigines, and common citizens.

Regarding commercial competitiveness, the national program gets help from MOEA: the Industrial Technology Department keeps doing researches and developing advanced learning technologies. Moreover, the Industrial Development Bureau draws many corporations in the e-learning industry to form a virtual science park in order to complete the e-learning supply-chain. Via aid of the supply-chain, e-learning industry can provide integrated and
complete services for the increase of commercial competence; commercial competence in various industries can improve through high quality e-learning products and services.

Regarding the researches in industry and academia, the national program assigns the Industrial Technology Department to learning technology related researches and developing patents; and the National Science Council not only offers funding for researchers in academic organizations, but also encourages the cooperation between academic researchers and industries in order to realize research results.

The national program has created 108 KPI items and divided them into 10 major dimensions in order to acquaint people and government with the quantitative outcomes of the national program. This paper shows some major quantitative KPI results and some qualitative cases about how this five-year national program promotes the public welfare, increases commercial competitiveness in both e-learning related and ordinary industries, and e-learning research results.

**Public Welfare Improvement**

As mentioned in Section 1, seven government agencies have participated in the operational program, "e-learning for everyone". Every agency has their targeted people. For example, Council of Labor Affairs mainly targets people who have been employed or unemployed and disabled workers; Ministry of Education mainly focuses on students in traditional education system and on job development people; Council for Culture Affairs mainly targets community reconstruction and cultural designers; Council of Indigenous Peoples mainly targets aboriginal people; Council for Hakka Affairs mainly targets Hakka people; Bureau of Health targets not only medical professionals but the public; and National Palace Museum also targets the public.

**Quantitative Results**

Two KPI dimensions, public promotion and e-learning, are used to represent "e-learning for everyone" operational program. The public promotion dimension demonstrates how many people might know of the national program for e-learning in Taiwan and what e-learning really is. As we can see from Table 2, in the five years the national program has held 378 public activities and the total participants add up to more than two millions (2,474,962).

The age range of participants falls mainly between 12 to 50 years old, and curiosity might be the reason for more participants in the first year. In the third year, 2005, public promotion project was held by the Taipei Computer Association, which is the most important computer association in Taiwan and had held the Computex exhibition several times. The exhibition has involved project presentation, governmental demonstration, and product marketing in attracting people's attentions. Thus the participant number was raised so much in 2005. In the following year, in 2006, the participant number dropped again.

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newspaper</td>
<td>58</td>
<td>121</td>
<td>147</td>
<td>303</td>
<td>323</td>
</tr>
<tr>
<td>Public Activities</td>
<td>41</td>
<td>28</td>
<td>87</td>
<td>130</td>
<td>92</td>
</tr>
<tr>
<td>Participants</td>
<td>23,448</td>
<td>205,572</td>
<td>1,950,332</td>
<td>203,197</td>
<td>92,413</td>
</tr>
</tbody>
</table>

The e-learning dimension shows how many e-learning websites have already been built (and most of them were free); how many people have been visited; how many courses have been developed; and moreover, how many items have been constructed. For the five years, 45,507 e-learning courses have been developed in over hundred e-learning websites, listed as table 3. Moreover, these e-learning courses also attract nearly eighteen millions (17,450,981) visiting counts to those e-learning websites.

While promoting the e-learning courses, the strategy is to acquaint people with the best practices. There are hundreds of thousands of mid- and small-scale enterprises, and correspondent industrial associations in Taiwan. The top one thousand enterprises have been selected as the promotion targets and the national program has helped them with in-house e-training materials. The national program sets up many courses to establish the industrial learning network for
business purposes; hence, there are 85,211.5 hours of e-learning materials in 2004 consequently. In 2005, the national program spreads the experiences to various industries suchlike fishing industry, farming industry, and manufacturing industry. Thus the e-learning material hours in 2005 dropped down hugely but the course number still increased from 2004’s 8,777 to 2005’s 20,431.

Table 3: KPI (e-Learning)

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Websites</td>
<td>31</td>
<td>58</td>
<td>43</td>
<td>53</td>
<td>120</td>
</tr>
<tr>
<td>Visitors</td>
<td>220,930</td>
<td>3,487,451</td>
<td>5,675,808</td>
<td>3,883,249</td>
<td>4,183,543</td>
</tr>
<tr>
<td>Courses</td>
<td>153</td>
<td>8,777</td>
<td>20,431</td>
<td>307</td>
<td>15,839</td>
</tr>
<tr>
<td>Hours</td>
<td>320</td>
<td>85,211.5</td>
<td>16,150</td>
<td>9,079.67</td>
<td>82,566.2</td>
</tr>
<tr>
<td>Items+</td>
<td>1,032</td>
<td>2,634</td>
<td>2,006</td>
<td>213,821</td>
<td>110,647</td>
</tr>
</tbody>
</table>

*total e-learning course hours
+in item bank

In the forth year, 2006, the national program focused on the assessments and the development of item bank. The assessment was used to improve the quality of e-learning course and to be the proof of offering people certifications. That’s why items of 2006 exceeded year 2005 by a hundred times.

In details, Council of Labor Affairs has developed 145 courses in five categories; Ministry of Education has developed 54 courses (equals to 364 hours); Council of Culture Affairs has developed 102 courses in ten categories; Council of Hakka Affairs has developed 72 courses (equals to 432.5 hours) in seven categories and 12 hours of children e-learning courses; Council of Indigenous Peoples has developed 56 courses; Bureau of Health has developed 76 courses (equals to 171 hours) for introducing ten major diseases in Taiwan to medical professions and the public; and National Palace Museum has developed 63 courses in seven categories and there are 15 courses in English.

Council of Labor Affairs announces an online testing system for professional certificates, people can get a better perception of their capability through the online simulate test. The online testing system has attracted more than 2.38 million visitors up to Sep. 2007. Council of Hakka Affairs builds Hakka language ability exam and related certificates for Hakka people. Council of Indigenous Peoples also sets up exams for language skills and related certificates of aboriginal languages, including forty aboriginal dialects in Taiwan. Furthermore, Bureau of Health has issued 2,909 medical training certificates, including training hour certificates to the publics and the civil servants (1,097 certificates), and the continuous education certificates to medical professionals (1812 certificates).

Qualitative Cases

Council of Labor Affairs develops e-learning courses and special e-learning website for disabled people. These courses help disabled people learn working skills and improve their personal competitiveness. Ministry of Education cultivates elementary and secondary school teachers in godforsaken places through e-learning courses. Bureau of Health helps the general public, no matter where they are living, getting preventive care and self care knowledge. National Palace Museum has more than 650,000 artifacts covering 7,000 years of Chinese history and has developed a series of e-learning courses in multiple languages. The e-learning program of NPM offers the public opportunity to ‘touch’ Chinese arts and culture virtually, furthermore, to learn the related knowledge about the collections online.

Furthermore, there are two cases that can represent the qualitative results in the public welfare, Mr. Arnold (nickname) and Mr. Wu's cases. Arnold is a young person who does not have high school degree and Mr. Wu is a salesman currently working in a frozen fish company, Gallant Ocean Group. They all get successful experiences via the national program.

Arnold is a person who used to do nothing but hanging out around the neighborhood; he doesn't have high school degree. He found the Cultural Affairs School of E-learning (CASE) when browsing the web pages. Arnold was attracted by those blended courses provided by Council for Cultural Affairs and decided to apply the online school. His semester project was saving a hydrographic station during Japanese colonial period. After Arnold studied in CASE online school, he became a community reconstruction designer and had many cases in his hands.
Mr. Wu was a book salesman before he joined Gallant Ocean Group. He didn't have any knowledge about frozen fish. Gallant Ocean Group built an e-learning environment for its employees. A salesman like Mr. Wu needed to learn knowledge before they visited their customers. Wu learned a lot from the e-learning courses, and he did so well that his customers thought he had been working in frozen fish business in his whole life. Mr. Wu became the number one salesman three months later after he joined Gallant Ocean Group.

**Increasing Business Competitiveness**

Figure 1 represents the market size of Taiwan e-Learning industry from 2002 to 2006. As we can see from Figure 1, before the national program for e-learning started, the market size was only 22.64 millions USD. Although the market size was growing steadily, after the national program started from 2003 the market size grew extremely fast.

![Figure 1. Market Size of Taiwan e-Learning Industry (source: ELNP progress report 2006). (note: the national program started from Jan. 2003)](image)

The Economist Intelligence Unit and IBM also conducted an e-Learning readiness white paper in 2004, in this report Taiwan sat on the 16th around world among 60 countries and was in 3rd place of Asia just behind the South Korea and Singapore (Economist Intelligence Unit & IBM, 2004). What’s notable is that e-Learning readiness of Taiwan in the business part was in 9th place of the world.

**Quantitative Results**

There are two KPI dimensions related to e-learning industry, professionals and economics. The industrial part of professionals lists how many multimedia instructors have been trained by the national program; how many instruction designers have been trained; how many multimedia developers have been trained; and how many IT professionals have been trained. Table 4 indicates that there are more than a thousand (1,142) instruction designers have been trained. In general, an e-learning course development team always has one instruction designer, two multimedia developers, and four to five IT professionals.

Course Planners are the first line personnel. The national program set up the curriculum in 2004 particularly to boost the development of e-learning. The number of the course planners in 2004 was increased to 209, but then decreased
gradually in the following years. The number of cultivated instructional designers was similar to the course planners number. In 2006, the instructional designers added up to 542 because the National Palace Museum needed hundreds of instructional designers to help them transfer digital archives into e-learning courses.

In the third year, 2005, the national program focused on having more multimedia developers because e-learning courses and instructional designing require multimedia development. There were many e-learning professionals learned in the two skills mentioned previously in first year and in the multimedia development skill in second year. The number of cultivated IT professionals in the forth year dropped down immensely because the IT talent training program had come to a halt.

The KPI items in economical dimension represent how many e-learning ISPs have been assisted by the national program and how much expenditure have been spent in applying e-learning. Regarding the expenses in Table 5, in the beginning, most enterprises kept a wait-and-see stand. One year later, some successful examples urged them to invest in their in-house e-learning training materials. Commercial expense for e-learning in 2004 was almost ten times than 2003. However, in the third year, 2005, most enterprises which wanted to invest in e-learning had already got involved, hence, the expenditure dropped down from 37.23 to 1.76. The phenomenon was also affected by the government awards. In 2004, the national program launched a three-stage reward mechanism and conferred the reward to the enterprises which were eligible for the reward requirements. In 2005 and 2006, the reward scale reduced and might affect the investing willingness in e-learning among enterprises.

Before the national program started, the e-learning market size was small; the quality of e-learning materials and services was not good enough; only one enterprise was involved in developing e-learning contents, and yet the quality of contents could not reach the standard of international market; there was no e-learning portal; there was no standards for either e-learning platform or e-learning contents; and there was no e-learning centers.

Now, the e-learning market size is around 314 millions USD; there are eighty-nine e-learning vendors certified for the quality of their e-learning contents; many e-learning industries have started to get orders all over the world, including the US, Japan, Korea, Thailand, India etc.; there are five e-learning portals with more than 3,500 courses and more than 320 thousand members; the interoperability is greater than 85% with SCORM standard; and there are 86 e-learning centers now.

**Qualitative Cases**

Due to the promotion of national program, industries have shifted from the ignorant to positive executants. Moreover, the decision process has also been shortened from "six moths" to "three months", which means applying e-learning to business can gain more profits. Taiwan is also the first nation which got ADL SCORM certificate and ASTD ECC quality certificate.

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multimedia Instructors</td>
<td>3</td>
<td>68</td>
<td>38</td>
<td>206</td>
<td>442</td>
</tr>
<tr>
<td>Course Planners</td>
<td>31</td>
<td>209</td>
<td>83</td>
<td>137</td>
<td>113</td>
</tr>
<tr>
<td>Instruction Designers</td>
<td>61</td>
<td>168</td>
<td>26</td>
<td>542</td>
<td>345</td>
</tr>
<tr>
<td>Multimedia Developers</td>
<td>32</td>
<td>51</td>
<td>354</td>
<td>97</td>
<td>70</td>
</tr>
<tr>
<td>IT Professionals</td>
<td>2,025</td>
<td>2,047</td>
<td>2,671</td>
<td>31</td>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assisting*</td>
<td>29</td>
<td>47</td>
<td>55</td>
<td>49</td>
<td>27</td>
</tr>
<tr>
<td>Business Expenses$</td>
<td>6.22</td>
<td>37.23</td>
<td>1.76</td>
<td>7.93</td>
<td>1.75</td>
</tr>
<tr>
<td>e-learning Market Size#</td>
<td>67.91</td>
<td>121.64</td>
<td>196.97</td>
<td>314.4</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*assist how many e-learning ISPs  
$in Million USD  
#data comes from Market Intelligence Center (MIC) of Institute for Information Industry, Taiwan
Following are two successful cases from the national program. The first case is Family Mart Co., Ltd. and the second case is TransAsia Airways. Family Mart, a chain convenience store, was introduced into Taiwan from Japan in 1988 and now has more than 2000 stores in Taiwan. TransAsia was established in 1951 and wanted to become the best domestic airline as well as a globally renowned airline in the Asia-Pacific region.

Family Mart has so many stores in Taiwan. Employees in the chain stores are mostly students doing part time, so the corporation usually invests much expenditure in personnel training. Besides, in order to compete with other similar chain stores in Taiwan, Family Mart has to come up with new lunch meals every two to four weeks. The problem is how to get all employees well-trained before a new lunch meal is releasing.

Family Mart took e-learning as a good methodology to provide employees fast and cheap on-job-training (OJT). Therefore, since 2003 Family Mart developed e-learning platform and courses for fast delivery to employees. Now whenever a new promotion is released, there’s no need to summon all managers or store representatives for conference. Family Mart delivers related e-learning materials directly to every store and saves a lot of budget for the company and a lot of traveling time for employees. According to Family Mart's estimation the opportunity cost of on-job-training was saving approximately more than 39 thousands USD in total (around 50% cut-off of budget) yearly as Table 10 shows.

<table>
<thead>
<tr>
<th>Item</th>
<th>Opportunity cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal Personnel</td>
<td>Save approximate 9.15/yr</td>
</tr>
<tr>
<td>Staffs</td>
<td>Save approximate 7.32/yr</td>
</tr>
<tr>
<td>Part-time students</td>
<td>Save approximate 22.86/yr</td>
</tr>
<tr>
<td>Total</td>
<td>Save approximate 39.33/yr</td>
</tr>
</tbody>
</table>

The number of passengers traveling by air has recently fallen due to the increasing price of gas and the new high-speed railway. To increase competition and adapt to the market environment, Porter (1997) recommended that corporations should first reduce cost and provide various services. Regular training requires a large expense for an airline. Therefore, to reduce the cost of training, TransAsia decided in 2004 to replace traditional training with e-learning.

TransAsia developed 60 e-learning courses, including 36 courses (lasting about 80 hours) produced in-house. According to TransAsia's estimation, it might save up to 800 thousand USD in total yearly as Table 10 shows (Chuang et al., 2008).

<table>
<thead>
<tr>
<th>Item</th>
<th>Opportunity cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight Crews' Salary</td>
<td>Save approximate 340/yr (save training time, increase flying time)</td>
</tr>
<tr>
<td>Operation Earnings</td>
<td>Save approximate 840/yr (will be able to have additional flights)</td>
</tr>
<tr>
<td>Lecturers' Expenses</td>
<td>Save approximate 20/yr (do not need to spend money for lecturers)</td>
</tr>
<tr>
<td>e-learning Expenses</td>
<td>Spend approximate 400/yr (course development, infrastructure construction, learning management system, e-learning website etc.)</td>
</tr>
<tr>
<td>Total</td>
<td>Save approximate 800/yr</td>
</tr>
</tbody>
</table>

TransAsia not only received the government reward (approximately sixty thousand USD), but was also named a winner (bronze medal) in the Innovative Technology Category in the 2005 Brandon Hall Excellence in Learning Awards.
Stimulating Researches

As we all know, promoting e-learning industry requires high quality e-learning products. There’s always know-how for high quality products. At this moment, the academic researching outputs are taking place. Three KPI dimensions here are used to evaluate the academic outcomes of the national program, which are academic dimension, international dimension, and professional dimension. Moreover, this section also shows evidences about how many patents and innovations have been applied, received, and made by the national program; how the program helps corporations with technological transferring from MOEA; and how much that the government has gotten back from industry while putting the investment into the national program.

Quantitative Results

The academic dimension shows how many SSCI/SCI EI journal papers and conference papers have been granted by National Science Council and published; how many industry cooperative projects have been initiated; how many professors and degree holders have joined NSC e-learning research projects; and moreover, how many future researchers and/or developers, that is, Ph. D students and Master students, have been involved in NSC research projects. Table 8 shows there are 355 SSCI/SCI/EI journal papers, 365 national journal papers and that 1,249 conference papers have been published already; and, there are more than 220 professors and Ph. D holders run NSC e-learning research projects this year.

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCI/SCI/EI</td>
<td>32</td>
<td>78</td>
<td>88</td>
<td>73</td>
<td>84</td>
</tr>
<tr>
<td>Nat’l Journal</td>
<td>75</td>
<td>63</td>
<td>70</td>
<td>89</td>
<td>68</td>
</tr>
<tr>
<td>Int’l Conf. Paper</td>
<td>85</td>
<td>133</td>
<td>75</td>
<td>189</td>
<td>160</td>
</tr>
<tr>
<td>Nat’l Conf. Paper</td>
<td>114</td>
<td>138</td>
<td>145</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>Industry cooperation</td>
<td>9</td>
<td>18</td>
<td>21</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>NSC projects</td>
<td>69</td>
<td>88</td>
<td>77</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>NSC project chairs</td>
<td>234</td>
<td>209</td>
<td>216</td>
<td>176</td>
<td>195</td>
</tr>
<tr>
<td>Other Ph.D holders</td>
<td>41</td>
<td>26</td>
<td>45</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>Ph.D students</td>
<td>60</td>
<td>110</td>
<td>108</td>
<td>85</td>
<td>74</td>
</tr>
<tr>
<td>Master students</td>
<td>303</td>
<td>401</td>
<td>411</td>
<td>215</td>
<td>190</td>
</tr>
<tr>
<td>RA – Master</td>
<td>64</td>
<td>84</td>
<td>70</td>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>RA - others</td>
<td>213</td>
<td>183</td>
<td>226</td>
<td>81</td>
<td>104</td>
</tr>
</tbody>
</table>

The international dimension demonstrates how many important international e-learning conferences have been held in Taiwan; how many e-learning related experts, scholars, and organizations have been invited to Taiwan for discussion and giving talks; and also, how many e-learning international activities such as international conferences that Taiwan government has funded. Table 9 shows that Taiwan held IEEE Wireless and Mobile Technologies in Education (WMTE) in 2004, Computer Supported Collaborative Learning (CSCL), IEEE International Conference on Advanced Learning Technologies (ICALT) in 2005, Intelligent Tutoring System (ITS), and IEEE System, Man, and Cybernetics (SMC) in 2006 for the last 3 years.

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int’l Conf (to hold)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Guest Inviting</td>
<td>41</td>
<td>17</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Participating Int’l Activities</td>
<td>76</td>
<td>69</td>
<td>12</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

*IEEE WMTE, CSCL, IEEE ICALT, ITS, IEEE SMC, APEC
*including experts/scholars/organizations
Following the academic dimension, the professional dimension could be divided into academia and industry as listed in Table 4 and Table 10. The professional dimension lists how many research assistants have been hired for the national program; how many students have gotten their master’s degree; and, how many students have gotten their PhD degree. Table 6 shows that 1,778 students have got their master’s degree in the first phase (2003-2007) and 789 students their PhD degree.

Table 10: Research Results KPI (Professionals – academia part)

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Assistants</td>
<td>362</td>
<td>532</td>
<td>921</td>
<td>772</td>
<td>577</td>
</tr>
<tr>
<td>Masters</td>
<td>269</td>
<td>430</td>
<td>512</td>
<td>528</td>
<td>389</td>
</tr>
<tr>
<td>PhDs</td>
<td>153</td>
<td>201</td>
<td>210</td>
<td>102</td>
<td>79</td>
</tr>
</tbody>
</table>

*research assistant number might be overlapped in different years

Industrial Technology Department of MOEA takes charge of the advanced learning technologies researching and development. The task involves developing new technologies and innovative ideas. Table 11 shows that how many innovations have got patents; how many patents have been received (as we all know that the reviewing process for patent usually takes one to two years); how many patents have been taken into production; how many corporations have been asked to transfer the invented technologies; and how much they have been paid by the government for transferring these technologies.

Table 11 shows that 75 innovations around the world have got patents, and 19 patents have been received. In the technology transfer statistics, 70 corporations have been asked for transferring 78 technologies from the government and have paid more than 1,568 thousand USD for the technologies. The reward of technology transfer has only been 1,100 thousands USD in the past four and half years, to the roughly budget of IDB MOEA, 1,600 thousand USD each year. Although the reward is not much, it is still growing up according to Table 11. And also, the purpose that Taiwan government proposes the five-year national program for e-learning is to stimulate and establish the e-learning industry rather than gain profits.

Table 11: Research Results KPI (Technology)

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent (applying)</td>
<td>19</td>
<td>19</td>
<td>11</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Patent (received)</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Patent (products)</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Innovation(^1)</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Technology Transfer</td>
<td>13</td>
<td>16</td>
<td>23</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Transfer Business(^2)</td>
<td>5</td>
<td>13</td>
<td>23</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Transfer Money(^3)</td>
<td>177.43</td>
<td>270.03</td>
<td>511.63</td>
<td>220.87</td>
<td>388.55</td>
</tr>
</tbody>
</table>

\(^1\) e-learning products according to patents
\(^2\) including innovative systems/tools/modules
\(^3\) how many businesses asked for technology transfer
\(^4\) in Thousand USD

**Qualitative Results**

Taiwan has held many important academic activities in the past four years, for example, IEEE Workshop on Wireless and Mobile Technologies in E-learning 2004, Computer Supported Collaborative Learning Conference 2005, IEEE International Conference on Advanced Learning Technologies 2005, and International Conference on Intelligent Tutoring Systems 2006. Taiwan also cooperated with the Advanced Distributed Learning Co-Laboratory (ADL Co-Lab) hosted International Plugfest II and held the 2006 International Conference on SCORM 2004 in conjunction with International Plugfest II.
Beside the international academic activities, Latchem (2006), the editor of British Journal of Educational Technology, did a quantitative survey on the publications in the past five years. The results showed that Taiwan was ranked the 4th for the publications from 2000 to 2005 in the Journal.

The national program also investigated in academic publication status in Taiwan in the following six SSCI journals. The data was queried from database in ISI Web of Science, from Jan. 2001 to Sep. 2006. The six journals are:

- Computers & Education (C&E)
- Journal of Computer Assisted Learning (JCAL)
- Educational Technology & Society (ET&S)
- ETR&D-Educational Technology Research and Development (ETR&D)
- Innovations in Education and Teaching International (IETI)
- British Journal of Educational Technology (BJET)

**Table 12:** Paper publication amounts (from ISI Web of Science, 2000.1.1 to 2007.12.17, ranked by country)

<table>
<thead>
<tr>
<th>Journal name</th>
<th>No.1</th>
<th>No.2</th>
<th>No.3</th>
<th>No.4</th>
<th>No.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&amp;E</td>
<td>England(115)</td>
<td>USA(86)</td>
<td><strong>Taiwan(52)</strong></td>
<td>Australia(27)</td>
<td>Netherlands(21)</td>
</tr>
<tr>
<td>JCAL</td>
<td>England(99)</td>
<td><strong>Taiwan(52)</strong></td>
<td>USA(30)</td>
<td>Netherlands(21)</td>
<td>Australia(17)</td>
</tr>
<tr>
<td>ET&amp;S</td>
<td>USA(75)</td>
<td>England(30)</td>
<td><strong>Taiwan(27)</strong></td>
<td>Greece(24)</td>
<td>Canada(22)</td>
</tr>
<tr>
<td>ETR&amp;D*</td>
<td>USA(176)</td>
<td>Netherlands(21)</td>
<td>Canada(8)</td>
<td>Australia(7)</td>
<td>South Korea(7)</td>
</tr>
<tr>
<td>IETI</td>
<td>England(101)</td>
<td>Australia(20)</td>
<td><strong>Taiwan(16)</strong></td>
<td>Scotland(12)</td>
<td>USA(12)</td>
</tr>
<tr>
<td>BJET†</td>
<td>England(207)</td>
<td>USA(83)</td>
<td>Australia(70)</td>
<td>Scotland(53)</td>
<td>India(34)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>England(553)</strong></td>
<td><strong>USA(462)</strong></td>
<td><strong>Taiwan(178)</strong></td>
<td><strong>Australia(158)</strong></td>
<td><strong>Netherlands(121)</strong></td>
</tr>
</tbody>
</table>

*Taiwan has one paper, No. 13
†Taiwan has 30 papers, No. 7

**Table 13:** Paper citation amounts (from ISI Web of Science, 2000.1.1 to 2007.12.17, ranked by country)

<table>
<thead>
<tr>
<th>Journal name</th>
<th>No.1</th>
<th>No.2</th>
<th>No.3</th>
<th>No.4</th>
<th>No.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&amp;E</td>
<td>USA(86)</td>
<td>187 times</td>
<td>England(115)</td>
<td>144 times</td>
<td><strong>Taiwan(52)</strong></td>
</tr>
<tr>
<td>JCAL</td>
<td>England(99)</td>
<td>164 times</td>
<td><strong>Taiwan(52)</strong></td>
<td>135 times</td>
<td>USA(30)</td>
</tr>
<tr>
<td>ET&amp;S</td>
<td>Netherlands(20)</td>
<td>39 times</td>
<td>England(30)</td>
<td>28 times</td>
<td>Greece(24)</td>
</tr>
<tr>
<td>ETR&amp;D*</td>
<td>USA(176)</td>
<td>381 times</td>
<td>Netherlands(21)</td>
<td>72 times</td>
<td>Australia(7)</td>
</tr>
<tr>
<td>IETI</td>
<td>England(101)</td>
<td>44 times</td>
<td><strong>Taiwan(16)</strong></td>
<td>24 times</td>
<td>Australia(20)</td>
</tr>
<tr>
<td>BJET†</td>
<td>England(207)</td>
<td>225 times</td>
<td>USA(83)</td>
<td>73 times</td>
<td>Australia(70)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>USA(462)</td>
<td>729 times</td>
<td>England(553)</td>
<td>652 times</td>
<td><strong>Taiwan(178)</strong></td>
</tr>
</tbody>
</table>

*Taiwan’s one paper has not been cited yet

**Table 14:** numbers of academic research papers in Taiwan in the six journals (from 2000.1.1 to 2007.12.17)

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJET</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td>C&amp;E</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>17</td>
<td>55</td>
</tr>
<tr>
<td>ET&amp;S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>ETR&amp;D</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>IETI</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>16</td>
</tr>
</tbody>
</table>
From Table 14, the academic papers published in the six journals have increased gradually. Although publication number has increased, the researches mainly focus on the system development and are limited to the laboratory and/or specific experimental environment. Therefore, the national program should focus on the learning environment and exam designs in practicality, encourage the long term research plans, create more opportunities for testing new systems in real environment, and encourage the cooperative research projects cross different domains or countries.

**Conclusions**

Taiwan government has set up a five-year national program for e-learning since 2003. Benefit that the industries have gained from the national program is obvious. And as showed in Figure 1, the e-learning market keeps growing amazingly. The national program not only benefits industries but also the public. As we can see that 36,187 e-learning courses have had been developed and attracted over eighteen millions visiting counts listed on Table 3. Meanwhile, two cases have proved how the national program can affect the public. Furthermore, the IDB of MOEA and the National Science Council have developed advanced learning technologies; brought researchers and enterprises into cooperation; turned research results into products and patent applicable; and transfer new technologies and innovations to e-learning vendors in order to increase their competitiveness and quality of e-learning products.

Some bottlenecks were encountered in the past five years, including the collaboration issues between different participant agencies and divisions each agency holds. For examples, CLA may want to emphasize on how e-learning supports employee to update working knowledge and unemployed people to have required working skills in order to get jobs; NPM may want to focus on how e-learning attracts visitors and teaching them knowledge of Chinese culture; and, MOEA may want to know how e-learning can help improve commercial competitiveness. Healthy collaboration among agencies is a huge issue for the national program. Taiwan government demands different participating agencies to work together and/or support each other. For example, CLA and MOEA construct some e-learning courses training the unemployed and delivering working knowledge, furthermore, MOE, CLA, and MOEA work together to build a certificate for technicians and designers of e-learning, the technicians and designers can then get into businesses and help the businesses apply e-learning into the organizations.

From 2008 to 2012, Taiwan government has set up another program for e-learning, Taiwan e-Learning and Digital Archives Program (TELDAP). The new program includes eight projects: Taiwan Digital Archives Expansion Project, Research and Development of Digital Archive and e-Learning Technologies Project, Core Platforms for Digital Contents Project, The Project of Academic and Social Promotion and Applications for Digital Archives and e-Learning Project, Industrial Development and Promotion of Digital Archives and e-Learning Project, ICT in Education and e-Learning Project, Chinese Language e-Learning Project, International Collaboration and Promotion of Taiwan Digital Archives and e-Learning Project. National Program for e-Learning works with National Digital Archives Program to popularize the knowledge in the archive; head to a knowledgeable society; nourish e-learning industries and develop Chinese e-learning system in practical aspect; deepen the influence both on formal education and lifelong learning; and create sustaining benefit as improving national competitiveness.

We expect Digital Archive and National Digital Archives Program to expand its influence on national human resources and art development; enable the ample of Taiwanese culture establish its subject position, entering the era of globalization and knowledgeable society synchronously with the developed countries in American and Europe; by digital learning, to develop national population quality and cultivate talents and establish digital learning industry, providing digital teaching services and products to the world.

**References**


An Evaluation of the ELNP e-Learning Quality Assurance Program: Perspectives of Gap Analysis and Innovation Diffusion

Ming-Puu Chen
Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan // mpchen@ntnu.edu.tw

ABSTRACT

The purpose of this study was to examine the appropriateness of a nationwide quality assurance framework for e-learning from participants’ perspectives. Two types of quality evaluation programs were examined in this study, including the e-Learning Service Certification program (eLSC) and the e-Learning Courseware Certification program (eLCC). Gap analysis and the innovation attributes were employed to examine participants’ perception gaps and attitudes toward the quality framework. The analysis showed that the quality framework obtained adequate validity and reliability. Gap analysis revealed that both the eLSC and eLCC participants perceived the quality of their e-learning overly. The attitude analysis revealed that a) for eLSC, the positive and increased observability was inferred to enhance participants’ adoption of eLSC, b) for eLCC, the positive attributes of observability, relative advantage and compatibility were inferred to facilitate and sustain the adoption of eLCC, c) the decreased attributes of trialability/complexity and relative advantage of eLCC indicated the necessity for further improvement. Crystal-clear descriptions and examples of quality criteria are suggested to make the quality assurance framework more amiable and easy of access to e-learning developers and organizations.

Keywords

E-learning, Quality assurance, Courseware, Innovation diffusion

Introduction

Recently, the application of the Internet and multi-media technologies to learning has transformed the traditional face-to-face learning into a new and pervasive variety of “anytime, anywhere, and anyone” e-learning. In responding to the rapid growth of e-learning applications, the “National Science and Technology Program for e-Learning (ELNP)” was launched by the National Science Council of the Executive Yuan of Taiwan with the aims to a) upgrade Taiwan's overall competitiveness in the era of knowledge economy, b) stimulate the development of e-learning related industries, and c) bring about new waves of academic research (ELNP, 2003). From the viewpoint of national development, bridging the digital divide between rural and urban areas has become an important issue. It is, therefore, imperative for the government to take the lead in building an affordable, effective, and easy-to-access e-learning environment, especially for the remote regions of the country. From the human resource viewpoint, the needs to train a great number of internationally competitive talents in order to compensate for the country’s limited space and natural resources are essential in maintaining its competitiveness in the knowledge economy era. Whereas quality e-learning can provide efficient and effective digital learning experiences to the learners through web-based interactive methods, it also plays an important role in improving the quality of human resources. ELNP is an integral part of the nation’s efforts to lead Taiwan into the knowledge economy century under the Council for Economic Planning and Development’s national development plan, “Challenge 2008”. The 5-year e-learning development program aimed to build a quality e-learning environment for education and training, with the hope to enhance the quality of human resources, bridge the digital divide between rural and urban regions in the country, and bring about better competitiveness of the nation in the knowledge economy century (ELNP, 2006). Therefore, it is hoped that ELNP will serve as a driving force for e-learning related industries and improve the nation’s overall competitiveness.

Although e-learning has become a popular way of learning, the educational effectiveness of the Internet-based e-learning is not automatically ensured through the use of the Internet and multi-media technologies. To enhance e-learning quality, the e-Learning Quality Assurance program (eLQA) was implemented in 2005 to recognize and promote quality e-learning under the ELNP national program. It was hoped that eLQA would serve as an accreditation service for e-learning providers targeting for corporate e-training as well as school e-learning. Therefore, eLQA provided the e-learning providers with opportunities to access to a reliable quality assurance system and to endorse the educational quality of e-learning for the certified e-learning organizations and courseware. It was also hoped that eLQA would provide trustful quality information for prospective e-learners, so they could
identify and select quality e-learning courseware and services confidently and truly benefited from e-learning. Although the introduction of eLQA has aroused general awareness toward quality issues in e-learning related industries in Taiwan, how participants perceive eLQA was not studied in a systematic manner yet. Therefore, the present study aimed to a) examine the validity and reliability of eLQA, b) explore the quality gaps between participants’ self-evaluation and the eLQA quality evaluation, and c) evaluate participants’ attitudes toward eLQA in terms of innovation attributes. Therefore, based on the results of the present study, eLQA can be further improved and become more amiable and easy of access to e-learning developers and organizations.

**Quality Measures for E-learning**

To ensure quality learning, many organizations and individuals have developed various quality measures for evaluating Internet-based e-learning (McLoughlin & Visser, 2003). Based on the perspective of needs assessment, Chapnick (2000) proposed an e-learning needs assessment model containing eight components to assess an e-learning organization’s “Psychological readiness”, “Sociological readiness”, “Environmental readiness”, “Human resource readiness”, “Financial readiness”, “Technology readiness”, “Equipment readiness”, and “Content readiness”. Focusing on an organization’s e-learning sustainability, Rosenberg (2000) suggested seven components to evaluate how an organization can sustain e-learning, including “Business readiness”, “Changing nature of learning and e-learning”, “Value of instruction and information”, “Role of change management”, “Reinvention of training organizations to support e-learning efforts”, “E-learning industry”, and “Personal commitment”. Similarly, Borotis and Poulammenakou (2004) defined seven components for evaluating e-learning readiness, including “Business readiness”, “Technology readiness”, “Content readiness”, “Training process readiness”, “Culture readiness”, “Human resources readiness”, and “Financial readiness”. Moreover, targeting on the e-learning readiness of a country, the Economist Intelligent Unit (2003) estimated the ability of 60 countries to produce, use and expand Internet-based learning in government, industry, education, and society by means of nearly 150 qualitative and quantitative criteria organized into four categories, including “Connectivity”, “Capability”, “Content” and “Culture” called 4Cs. “Connectivity” represents the quality and extent of the Internet infrastructure. “Capability” is the ability to deliver, consume, and develop e-learning in training and education. “Content” is the quality and pervasiveness of e-learning materials. Finally, “Culture” relates to behaviors, beliefs and institutions that support e-learning development.

In higher education e-learning, the Institute for Higher Education Policy (IHEP, 2000) identified 24 benchmarks considered essential to ensure excellence in Internet-based learning. These benchmarks were distilled from the most popular strategies employed by colleges and universities, and they were divided into seven categories of quality measures, including “Institutional support”, “Course development”, “Teaching and learning”, “Course structure”, “Student support”, “Faculty support”, and “Evaluation and assessment”. Likewise, the Western Cooperative for Educational Telecommunications (WCET) developed a set of “Principles of good practice for electronically offered higher education degree and certificate programs” that encompassed three categories of quality measures, “Curriculum and instruction”, “Institutional context and commitment”, and “Evaluation and assessment” (WCET, 2001, 2005). These principles have become a generally accepted basis for evaluating distance learning programs in the United States.

With respect to the quality of e-learning courses, the WebCT Exemplary Course Project annually recognizes e-learning courses that model best practices in “Course design”, “Interaction and collaboration”, “Assessment and evaluation”, “Meaningful technology use”, and “Learner support” (WebCT, 2006). Likewise, the annual Brandon Hall Research Excellence in Learning Awards employs seven criteria, including “Navigation”, “Content presentation”, “Use of media”, “Interactivity”, “Engagement”, “Support of objectives”, and “Overall”, to recognize the best custom-designed, self-paced, online, interactive courses related to workplace learning (Brandon Hall Research, 2008). Furthermore, the Quality Criteria for E-Learning Courseware designed by the American Society for Training and Development (ASTD) aims to raise the quality of asynchronous web-based and multimedia e-learning courseware in institutions and corporations through the E-learning Courseware Certification program (ECC). The 19 quality criteria of ASTD ECC evaluate the quality of “compatibility”, “interface”, “production quality”, and “instructional design” of an e-learning courseware (ASTD, 2002, 2006).

The growing emphases on quality measures among organizations and researchers revealed the importance of quality assurance for e-learning. In spite of the variety of quality measure for e-learning, there is consensus emerging as to
what constitutes quality e-learning (IHEP, 2000; McLoughlin & Visser, 2003). The readily available quality instruments can provide systematic frameworks for assessing e-learning, the results of the analysis, however, may vary from one instrument to another as well as from one organization to another due to the specific perspective and purpose of a quality instrument (Borotis & Poulymenakou, 2004). Therefore, an e-learning quality instrument should be utilized and interpreted thoughtfully and in accordance with what it aims to do. With the intention to support the continuous improvement of quality e-learning, the eLQA quality framework was developed to promote and encourage the e-learning industries to pursue high quality e-learning. Although the construct validity of the eLQA quality framework can be sustained by experts’ review and referred to the commonly recognized indexes, how the eLQA quality framework works is not assessed in a systematic manner, especially from the participants’ perspective.

The Innovation Diffusion Perspective

The quality of e-learning is a perception which needs to be verified. Although the eLQA quality framework aimed to serve as a common platform for ensuring quality e-learning in the country, the precondition of achieving the goal is based on e-learning developers’ positive attitudes toward and precise perception of the quality framework. Therefore, how e-learning participants perceive a specific quality measure needs to be studied in order to determine a best way to ensure its sustainable harvest. For improving program evaluation, Hubbard and Sandmann (2007) suggested that the empirically tested innovation diffusion theories, developed in the 1950s, have provided a popular framework to study and explain how new ideas are spread and adopted in a community. Based on Rogers’ perspective of innovation diffusion (Rogers, 2003), adopters of an e-learning quality measure will go through a five-stage innovation-decision process from knowledge of the quality measure to forming attitudes toward the quality measure, to a decision to adopt or reject the quality measure, to implementation and use of the new idea, and to confirmation of the adoption decision. Therefore, adopters’ attitudes toward the eLQA quality framework play an important role in making and sustaining the adoption decision, and inevitably affect the diffusion of the quality measure to the potential adopters.

Rogers (1995) concluded from studies and suggested that the way adopters perceive the attributes of an innovation is critical and these perceptions account for 49-86% of the variance in adoption or rejection of an innovation. The five most important attributes impacting on the rate of adoption of an innovation include the relative advantage, compatibility, complexity, observability, and trialability of an innovation (Rogers, 1995, 2003). **Relative advantage** is the degree to which an innovation is perceived as better than the one it supersedes, such as increased performance, reduced cost, or other superior factors. **Compatibility** involves the degree to which an innovation is perceived as being consistent with the existing values, past experiences and needs of potential adopters. **Complexity** involves the degree to which an innovation is perceived as difficult to understand and use from the potential adopters’ perspective. **Trialability** deals with the degree to which an innovation may be experimented with on a smaller and limited scale. Finally, **Observability** relates to the degree to which the results of the implemented practice are visible to others. Some practices are more observable than others and therefore might be adopted more quickly by potential users.

Moreover, Rogers (2003) noted that each construct of the innovation attributes is somewhat interrelated. Some studies (Goldman, 1994; Sonnenwald, Maglaughlin, & Whitton, 2001) concluded that the five constructs of the innovation attributes were distinct. Pankratz, Hallfors and Cho (2002), however, found three distinct constructs of advantage/compatibility, complexity and observability with the relative advantage and compatibility clustered into one construct. Due to the interrelated characteristics among these five constructs, Scheirer (1990) even suggested that the five perceived attributes can be combined into one index of favorability toward the adoption of an innovation. Therefore, based on the verified validity and appropriate accountability in explaining adopters’ perceptions of the adoption of an innovation, the constructs of innovation attributes are suitable for evaluating the diffusion of the quality framework. In the present study, the constructs of innovation attributes were employed to examine participants’ attitudes toward the adoption of eLQA.

The eLQA Quality Assurance Programs

In 2004, the eLQA quality framework was developed by means of a series of focus groups and with referring to previous quality measures, such as WCET (2001) and ASTD (2002), to ensure the construct validity of the quality
framework. The first version of eLQA quality assurance programs was implemented in January 2005 by the e-Learning Quality Certification Center (eLQC) which was setup and dedicated to providing quality certification services to e-learning organizations and e-learning courseware developers, especially for those stationed in the e-Learning Network Science Park. eLQA consists of two quality certification programs, the e-Learning Service Certification program (eLSC) and the e-Learning Courseware Certification program (eLCC), targeting on e-learning service providers and e-learning courseware developers, respectively. The eLSC quality evaluation assesses the quality of an e-learning organization from aspects of personnel, course and system. Meanwhile, the eLCC quality evaluation examines the educational quality of an e-learning courseware from aspects of content, navigation, instructional design and instructional media.

The e-Learning Service Certification Program

The development of eLSC quality evaluation was based on the perspective of “quality is how the organization goes about achieving its objectives” and aimed to recognize and promote quality e-learning services through a systematic process to evaluate the quality of e-learning service. As shown in Table 1, the quality framework of eLSC consists of three aspects of quality measures, including personnel, course and system. These three aspects can be further divided into eight quality criteria, including (1) learner support, (2) faculty support, (3) curriculum development, (4) instructional design, (5) instructional process, (6) organizational support, (7) technology and (8) assessment and evaluation. Each eLSC quality criterion comprises three to five sub-criteria with a total of 27 sub-criteria for the quality framework of eLSC.

Considering the various types and scopes of e-learning applications among organizations and companies, three types of e-learning service certification, including the Unit Certification, the Course Certification and the Curriculum Certification, were provided by ELNP to best fit the target audiences’ needs. The Unit Certification was targeted on organizations that do not need to apply whole scale instructional design in the development of e-learning courses. In general, this type of organizations is usually limited by time constraint and develops and delivers short and efficient e-learning just enough for internal needs. For instance, companies employed rapid e-learning methods to develop e-training to employees. In contrast, the Course Certification was targeted on organizations that always applied whole scale instructional design in the development of e-learning courses. This type of organizations develops and delivers self-directed e-learning to employees or to the clients. Finally, the Curriculum Certification was targeted on organizations that not only apply whole scale instructional design in the development of e-learning courses, but also provide certificates or degree programs to the clients, such as, training organizations and universities. As shown in Table 1, to apply for the Unit Certification, eLSC quality criteria 1, 2, 7 and 8 are required in the quality evaluation; to apply for the Course Certification, quality criteria 1, 2, 4, 5, 7 and 8 are needed; and to apply for the Curriculum Certification all the eight criteria are required.

Table 1. eLSC consists of three quality aspects, eight quality criteria and twenty-seven sub-criteria

<table>
<thead>
<tr>
<th>Quality aspects</th>
<th>Criteria</th>
<th>Number of sub-criteria</th>
<th>Unit Certification</th>
<th>Course Certification</th>
<th>Curriculum Certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnel</td>
<td>1. Learner support</td>
<td>3</td>
<td>■ Required</td>
<td>■ Required</td>
<td>■ Required</td>
</tr>
<tr>
<td></td>
<td>2. Faculty support</td>
<td>4</td>
<td>■ Required</td>
<td>■ Required</td>
<td>■ Required</td>
</tr>
<tr>
<td>Course</td>
<td>3. Curriculum development</td>
<td>3</td>
<td>□ Optional</td>
<td>□ Optional</td>
<td>■ Required</td>
</tr>
<tr>
<td></td>
<td>4. Instructional design</td>
<td>3</td>
<td>□ Optional</td>
<td>■ Required</td>
<td>■ Required</td>
</tr>
<tr>
<td></td>
<td>5. Instructional process</td>
<td>3</td>
<td>□ Optional</td>
<td>■ Required</td>
<td>■ Required</td>
</tr>
<tr>
<td>System</td>
<td>6. Organizational support</td>
<td>3</td>
<td>□ Optional</td>
<td>□ Optional</td>
<td>■ Required</td>
</tr>
<tr>
<td></td>
<td>7. Technology</td>
<td>5</td>
<td>■ Required</td>
<td>□ Optional</td>
<td>■ Required</td>
</tr>
<tr>
<td></td>
<td>8. Assessment and evaluation</td>
<td>3</td>
<td>■ Required</td>
<td>■ Required</td>
<td>■ Required</td>
</tr>
</tbody>
</table>

The eLSC quality evaluation employed a checklist-type evaluation form to identify the quality level of an organization’s e-learning service. For each sub-criterion, a 4-level checklist with quality levels of “AAA”, “AA”, “A”, and “Fail” was used to identify the quality level of an organization from a specific perspective. Due to the principle of “all criteria are required” of the eLSC quality evaluation, an applicant must pass all of the required sub-criteria of a specific type of certification, such as the Unit Certification, the Course Certification, or the Curriculum Certification, in order to be certified with the quality level of the lowest passing-level of this specific type of
certificate. For instance, if an applicant passes all of the six required sub-criteria of the Course Certification with a lowest quality level of “Level A”, then the applicant will be certified with a “Level A” Course Certificate. Furthermore, if a Course Certification applicant failed in any sub-criteria of the required “Course aspect” criteria, it will be downgraded to the Unit Certification evaluation. In other words, the lower type of eLSC certification will be applied automatically to the unqualified cases when it is applicable.

The e-Learning Courseware Certification Program

The eLCC quality evaluation aimed to recognize and promote well-designed, self-paced e-learning courseware through examining the educational quality of e-learning courseware from four quality aspects, including content, navigation, instructional design and instructional media. The quality aspects, quality criteria and objectives of quality aspects are shown in Table 2. The eLCC quality evaluation is composed of 15 quality criteria, which can be further identified as 8 required criteria and 7 optional criteria. eLCC scores each quality criterion with different weights to differentiate the importance of required and optional quality criteria. Each eLCC quality criterion comprises three checklist items to differentiate e-learning courseware into four levels of quality from a specific perspective. Accordingly, 0, 3, 5, or 7 points will be given for each required quality criterion and 0, 2, 4, or 6 points will be given for each optional criterion. Exceptionally, due to the importance of learning strategies, the optional quality criterion of “learning strategies” is scored as a required criterion. Therefore, as shown in Table 2, the eLCC quality evaluation measures the quality of e-learning courseware on a 100-point scale.

The eLCC quality evaluation employed individual reviews and a joint meeting of reviewers to determine the quality level of an e-learning courseware. The eLCC reviewers consisted of content experts, e-learning experts, academicians, instructional design practitioners and e-learning leaders in the industry. Individual reviews were conducted first by two randomly selected reviewers and one domain expert, and then a joint meeting of reviewers was held to make decisions of passing or failing for each eLCC quality criterion based on the results of individual evaluation. The qualified e-learning courseware was certified with “Level A”, “Level AA”, or “Level AAA” for passing all of the eight required criteria and with a total score equal to or higher than 60, 75 and 90 points, respectively.

<table>
<thead>
<tr>
<th>Quality aspects and criteria</th>
<th>Objectives of quality aspects</th>
<th>Points given by quality level</th>
<th>Sub-total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Content</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-1 Accuracy (r)</td>
<td>E-learning courseware should provide the learners with accurate, appropriately organized and clearly expressed content in order to facilitate the expected learning.</td>
<td>0, 3, 5, 7</td>
<td>21</td>
</tr>
<tr>
<td>1-2 Organization (r)</td>
<td></td>
<td>0, 3, 5, 7</td>
<td></td>
</tr>
<tr>
<td>1-3 Clarity (r)</td>
<td>expected learning.</td>
<td>0, 3, 5, 7</td>
<td></td>
</tr>
<tr>
<td><strong>2. Navigation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-1 Learning navigation (r)</td>
<td>E-learning courseware should provide the learners with navigational tools to facilitate smooth progress and effective management of learning.</td>
<td>0, 3, 5, 7</td>
<td>19</td>
</tr>
<tr>
<td>2-2 Operational helper (o)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td>2-3 Learner tracking (o)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td><strong>3. Instructional design</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-1 Goal and objectives (r)</td>
<td>E-learning courseware should provide the learners with well-designed learning activities, such as clearly expressed objectives and content, appropriate learning methods and strategies, and adequately designed practice, feedback and assessment, to facilitate learning interaction, comprehension and elaboration.</td>
<td>0, 3, 5, 7</td>
<td>41</td>
</tr>
<tr>
<td>3-2 Instructional presentation (r)</td>
<td></td>
<td>0, 3, 5, 7</td>
<td></td>
</tr>
<tr>
<td>3-3 Practice and feedback (r)</td>
<td></td>
<td>0, 3, 5, 7</td>
<td></td>
</tr>
<tr>
<td>3-4 Assessment (o)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td>3-5 Learning strategies (o)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td>3-6 Congruence (r)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td><strong>4. Instructional media</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-1 Media design (r)</td>
<td>E-learning courseware should employ well-designed instructional media to facilitate learning.</td>
<td>0, 3, 5, 7</td>
<td>19</td>
</tr>
<tr>
<td>4-2 Interface (o)</td>
<td>comprehension and sustain motivation for learning.</td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
<tr>
<td>4-3 Media elements (o)</td>
<td></td>
<td>0, 2, 4, 6</td>
<td></td>
</tr>
</tbody>
</table>

Note. (r): required criterion, (o): optional criterion, Total points = 100
Gap Analysis of eLSC and eLCC

The implementation of eLQA aimed to enhance the quality of e-learning service and courseware in the country. How precise the participants perceive the educational meanings of the quality criteria would affect the educational quality implemented in their e-learning products. Therefore, it was important to investigate whether any perception gap exists in the participants. Accordingly subsequent plans could be adopted to enhance learning quality and bridge the gap. Hence, the purpose of gap analysis was to examine whether gaps existed between the eLQA quality evaluation and applicants’ self-evaluation. For considering the congruence of quality evaluation standards, applicants evaluated by the 1.04 version of eLQA between July 2005 and June 2006 were selected and analyzed in the gap analysis. Totally, 28 eLSC cases and 37 eLCC cases were examined in the present study. Self-evaluation was a requirement for submitting an application for eLSC or eLCC certification. The eLQA quality evaluation was the final results of applicants’ eLSC or eLCC quality evaluation. Paired t-tests were employed to examine the difference between applicants’ self-evaluation and the eLQA evaluation, therefore possible gaps of participants’ perception of e-learning quality can be verified against the eLQA quality standards. The level of significance was set to .05 for all paired t-tests.

Gap Analysis of eLSC Cases

Twenty-eight eLSC cases were examined to identify gaps between the eLSC quality evaluation and applicants’ self-evaluation at the criterion level. For conducting quantitative analysis, the quality levels of “AAA”, “AA”, “A” and “Fail” of the eLSC cases were transformed to the values of 3, 2, 1 and 0, respectively, at the criterion level. Most of these applicants participated in the “Project of Promoting and Developing E-Learning Industries” conducted by the Industry Bureau of the Ministry of Economic Affairs, and being certified by eLQA was a requirement for winning awards. As shown in Table 3, there were 13 cases (46.4%) applied for the eLSC Course Certification and only 4 cases (14.2%) certified with the eLSC Course Certificate. Only 15 cases (53.6%) applied for the eLSC Unit Certification, but there were 24 case (85.7%) receiving the eLSC Unit Certificate. In other words, 9 cases applying for the Course Certification were downgraded and only certified with the Unit Certificate.

<table>
<thead>
<tr>
<th>Type of certificate</th>
<th>Number of cases</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit Certification</td>
<td>15</td>
<td>53.6%</td>
</tr>
<tr>
<td>Course Certification</td>
<td>13</td>
<td>46.4%</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>100.0%</td>
</tr>
<tr>
<td>Certificate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit Certificate, Level A</td>
<td>6</td>
<td>21.4%</td>
</tr>
<tr>
<td>Unit Certificate, Level AA</td>
<td>18</td>
<td>64.3%</td>
</tr>
<tr>
<td>Course Certificate, Level A</td>
<td>2</td>
<td>7.1%</td>
</tr>
<tr>
<td>Course Certificate, Level AA</td>
<td>2</td>
<td>7.1%</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

The construct validity of the eLSC quality evaluation was ensured by means of a series of focus groups during the development of the eLSC quality framework. The content validity coefficient of eLSC was .85 as measured by Kendall’s coefficient of concordance. The internal consistency reliability of the eLSC quality evaluation was .88 as measured by Cronbach’s α. The validity and reliability coefficients indicated that the eLSC quality evaluation obtained appropriate reliability and content validity.

Paired t-tests were conducted to examine whether gaps existed at the criterion level between the eLSC evaluation and applicants’ self-evaluation. The group means and summary of paired t-tests (self-evaluation – eLSC evaluation) by quality criterion are shown in Table 4. All of the means of the quality levels evaluated by the applicants were significantly higher than the eLSC evaluation. The results revealed that the eLSC applicants overestimated the quality levels of their e-learning in all quality criteria. The concordance of the lowest self-evaluation score (mean=1.87) and the lowest eLSC evaluation score (mean=1.00) assessed in the “Instructional design” criterion indicated that the eLSC participants were aware of the shortage in providing appropriate instructional design for supporting successful e-learning. The reasons why suitable efforts were not implemented in the instruction design for
the eLSC cases could not be answered in the present study. Moreover, whether the perception gaps were caused by participants’ misconceptions toward the quality measure, approaches of e-learning design, or other factors needs to be further studied.

Table 4. Summary of paired t-tests of eLSC cases by quality criterion

<table>
<thead>
<tr>
<th>Quality Criterion</th>
<th>Evaluation</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Learner support</td>
<td>Self-evaluation</td>
<td>2.45</td>
<td>.58</td>
<td>4.07*</td>
<td>27</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>2.04</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Faculty support</td>
<td>Self-evaluation</td>
<td>2.27</td>
<td>.60</td>
<td>3.90*</td>
<td>25</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>1.92</td>
<td>.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Instructional design</td>
<td>Self-evaluation</td>
<td>1.87</td>
<td>.45</td>
<td>4.81*</td>
<td>9</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>1.00</td>
<td>.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Instructional process</td>
<td>Self-evaluation</td>
<td>2.27</td>
<td>.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>1.60</td>
<td>.54</td>
<td>3.58*</td>
<td>9</td>
<td>.006</td>
</tr>
<tr>
<td>7. Technology</td>
<td>Self-evaluation</td>
<td>2.38</td>
<td>.52</td>
<td>2.88*</td>
<td>27</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>2.10</td>
<td>.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Assessment and evaluation</td>
<td>Self-evaluation</td>
<td>2.20</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eLSC evaluation</td>
<td>1.81</td>
<td>.45</td>
<td>3.46*</td>
<td>27</td>
<td>.002</td>
</tr>
</tbody>
</table>

Note. * p<.05, N=28

Gap Analysis of eLCC Cases

Thirty-seven e-learning courseware cases were examined to identify perception gaps between the eLCC quality evaluation and applicants’ self-evaluation at the quality aspect level. The distribution of the 37 eLCC cases was summarized in Table 5. For the field of use, 16 eLCC cases (43.2%) were developed for corporate training, 13 eLCC cases (35%) were developed for use in schools, and the other 8 cases (21.6%) were not limited to a specific field of use. As for the target audience, 25 cases (67.6%) were aimed at adult learning, 8 cases (21.6%) were targeted on school students, and the other 4 cases (10.8%) claimed to be suitable for audience of all ages. For the length of learning, about half of the eLSC cases (45.9%) were less than 2 hours of learning. Finally, as for the final results of the eLCC quality evaluation, only 2 cases (5.4%) were certified as eLCC Level AAA, 9 cases (24.3%) were certified as eLCC Level AA, 4 cases (10.8%) were certified as eLCC Level A, and the other 22 cases (59.5%) failed in the eLCC quality evaluation for not passing all of the 8 required criteria or with a total score less than 60 points.

Table 5. The distribution of eLCC cases by field of use, target audience, length of learning and quality level

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
<th>Number of cases</th>
<th>Percentage</th>
<th>Cum. percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field of use</td>
<td>Corporate training</td>
<td>16</td>
<td>43.2</td>
<td>43.2</td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>13</td>
<td>35.1</td>
<td>78.4</td>
</tr>
<tr>
<td></td>
<td>Not limited</td>
<td>8</td>
<td>21.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Target audience</td>
<td>Adult</td>
<td>25</td>
<td>67.6</td>
<td>67.6</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>8</td>
<td>21.6</td>
<td>89.2</td>
</tr>
<tr>
<td></td>
<td>Not limited</td>
<td>4</td>
<td>10.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Length of learning</td>
<td>16hr ~</td>
<td>6</td>
<td>16.3</td>
<td>16.3</td>
</tr>
<tr>
<td></td>
<td>6–16hr</td>
<td>7</td>
<td>18.9</td>
<td>35.2</td>
</tr>
<tr>
<td></td>
<td>2–6hr</td>
<td>7</td>
<td>18.9</td>
<td>54.1</td>
</tr>
<tr>
<td></td>
<td>0–2hr</td>
<td>17</td>
<td>45.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Certified quality</td>
<td>Level AAA</td>
<td>2</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>Level AA</td>
<td>9</td>
<td>24.3</td>
<td>29.7</td>
</tr>
<tr>
<td></td>
<td>Level A</td>
<td>4</td>
<td>10.8</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>Uncertified</td>
<td>22</td>
<td>59.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. N=37
The construct validity of the eLCC quality evaluation was ensured by means of a series of experts’ reviews, focus groups and pilot-tests during the development of the eLCC quality framework. The content validity coefficient of eLCC was .83 as measured by Kendall’s coefficient of concordance. The internal consistent reliability of the eLCC quality evaluation was .81 (Cronbach’s α) as measured on the 37 eLCC cases. It was concluded that the eLCC quality evaluation possessed adequate reliability and content validity.

The difference between the eLCC evaluation and applicants’ self-evaluation on each eLCC quality aspect was analyzed by means of paired t-tests. The mean scores and summary of the paired t-tests are shown in Table 6. The mean scores of the eLCC evaluation were significantly lower than the mean scores of applicants’ self-evaluation in all of the eLCC quality aspects. The results indicated that the e-learning courseware developers overestimated the educational quality of their e-learning courseware. Due to technical issues, cost, time, or other factors, the lowest scored-ratio among self-evaluation fell in the “Navigation” criterion (mean ratio=79%). Accordingly, the lowest scored-ratio among eLCC evaluation score fell in the “Navigation” criterion (mean ratio=60%) and revealed that the “learner tracking” sub-criterion was not fulfilled in most eLCC cases. Moreover, the largest perception gap (92% vs. 62%) was observed in the quality aspect of “Instructional design” and indicated that most of the eLCC applicants were over-confident with the instructional design features implemented in their e-learning products. The discrepancy could be derived from the developers’ lack of appropriate instructional design knowledge or due to the inadequate understanding toward the eLCC quality criteria. Further studies are suggested to explore the factors causing applicants’ perception gap in the instructional design aspect.

### Table 6. Group means and summary of paired t-tests of eLCC cases

<table>
<thead>
<tr>
<th>Quality aspect</th>
<th>Evaluation</th>
<th>Mean</th>
<th>Ratio</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>Self-evaluation</td>
<td>20.68</td>
<td>98%</td>
<td>1.11</td>
<td>6.36*</td>
<td>36</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>eLCC evaluation</td>
<td>16.84</td>
<td>80%</td>
<td>3.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navigation</td>
<td>Self-evaluation</td>
<td>15.00</td>
<td>79%</td>
<td>4.03</td>
<td>5.12*</td>
<td>36</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>eLCC evaluation</td>
<td>11.49</td>
<td>60%</td>
<td>4.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional design</td>
<td>Self-evaluation</td>
<td>37.81</td>
<td>92%</td>
<td>3.53</td>
<td>8.82*</td>
<td>36</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>eLCC evaluation</td>
<td>25.59</td>
<td>62%</td>
<td>8.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional media</td>
<td>Self-evaluation</td>
<td>17.97</td>
<td>86%</td>
<td>2.19</td>
<td>6.92*</td>
<td>36</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>eLCC evaluation</td>
<td>15.05</td>
<td>72%</td>
<td>2.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Self-evaluation</td>
<td>91.46</td>
<td>91%</td>
<td>10.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eLCC evaluation</td>
<td>68.97</td>
<td>69%</td>
<td>19.22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * p<.05, N=37

### Analysis of Participants’ Attitudes toward eLQA

The diffusion of the eLQA quality framework to e-learning developers and organizations cannot be accomplished by just a single event of announcing it. Instead, achieving a pervasive adoption of an innovation by the potential adopters relies on appropriate diffusion strategies and competitive attributes of the quality framework itself. The present study examined the diffusion of the eLQA quality framework from the perspective of perceived innovation attributes. Therefore, e-learning project managers in the 67 e-learning organizations stationed in the e-Learning Network Science Park were selected as the potential sample for data collection in the present study. The target population had involved in managing the quality of e-learning projects and was familiar with the quality assurance programs. Therefore, the target population is suitable for representing the adopters and prospective adopters of eLQA.

An on-line questionnaire was developed based on Rogers’ perspective of innovation attributes (Rogers, 2003) and conducted in February 2007 to investigate eLQA participants’ perception toward the implemented quality frameworks. As shown in Appendix A, the attitude questionnaire employed a 5-point Likert-type scale to collect participants’ perception of the trialability, complexity, observability, relative advantage and compatibility toward eLQA with 1 to 5 standing for “strongly disagree”, “disagree”, “neutral”, “agree” and “strongly agree”, respectively. There were three items for each component measure of the attitude questionnaire. Totally, the attitude questionnaire consisted of 15 items. For ensuring the validity, the survey instrument was developed with referring to previous
studies (Goldman, 1994; Hubbard & Sandmann, 2007; Pankratz, Hallfors, & Cho, 2002; Scheirer, 1990; Sonnenwald, Maglaughlin, & Whitten, 2001), reviewed and revised by peer experts, tested for validity and implemented in the present study. There were 71 responses received from 185 invitations with a response rate of 38.4%.

To assess the construct validity of the instrument, a principal component factor analysis was employed on participants’ perceptions of the innovation attributes toward eLQA. The mediocre KMO (.67) and significant Bartlett’s test of sphericity (Chi-square=2429.70, df=105, \(p<.001\)) indicated the appropriateness of conducting factor analysis. The summary of factor analysis is shown in Table 7, four factors, including trialability/complexity, observability, relative advantage and compatibility, emerged with eigenvalues greater than one based on Kaiser’s rule. Trialability and complexity were clustered as one combined construct, instead of two distinct constructs. The total variance explained by the four factors reached 89.53% of the variance. The overall reliability coefficient of the questionnaire was .91 as measured by Cronbach’s \(\alpha\), and the reliability coefficients of the component measures of trialability/complexity, observability, relative advantage and compatibility were .76, .79, .90 and .71, respectively, as measured by Cronbach’s \(\alpha\).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Trialability 2</td>
<td>0.92</td>
<td>-0.05</td>
<td>0.15</td>
<td>0.23</td>
</tr>
<tr>
<td>Trialability 3</td>
<td>0.85</td>
<td>0.26</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>Complexity 3</td>
<td>0.85</td>
<td>0.16</td>
<td>0.01</td>
<td>-0.04</td>
</tr>
<tr>
<td>Trialability 1</td>
<td>0.84</td>
<td>0.08</td>
<td>0.15</td>
<td>0.42</td>
</tr>
<tr>
<td>Complexity 2</td>
<td>0.81</td>
<td>0.47</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>Complexity 1</td>
<td>0.71</td>
<td>0.47</td>
<td>0.22</td>
<td>0.03</td>
</tr>
<tr>
<td>Observability 3</td>
<td>0.08</td>
<td>0.98</td>
<td>-0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Observability 2</td>
<td>0.19</td>
<td>0.94</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>Observability 1</td>
<td>0.46</td>
<td>0.84</td>
<td>0.11</td>
<td>-0.06</td>
</tr>
<tr>
<td>Relative Advantage 2</td>
<td>-0.05</td>
<td>-0.11</td>
<td>0.97</td>
<td>0.07</td>
</tr>
<tr>
<td>Relative Advantage 1</td>
<td>0.14</td>
<td>0.10</td>
<td>0.94</td>
<td>0.24</td>
</tr>
<tr>
<td>Relative Advantage 3</td>
<td>0.23</td>
<td>0.33</td>
<td>0.86</td>
<td>0.21</td>
</tr>
<tr>
<td>Compatibility 3</td>
<td>0.26</td>
<td>0.11</td>
<td>0.17</td>
<td>0.93</td>
</tr>
<tr>
<td>Compatibility 2</td>
<td>0.36</td>
<td>0.05</td>
<td>0.04</td>
<td>0.84</td>
</tr>
<tr>
<td>Compatibility 1</td>
<td>-0.13</td>
<td>-0.17</td>
<td>0.34</td>
<td>0.81</td>
</tr>
<tr>
<td>Cronbach’s (\alpha)</td>
<td>.76</td>
<td>.79</td>
<td>.90</td>
<td>.71</td>
</tr>
</tbody>
</table>

Note. 1. Extraction Method: Principal Component Analysis 2. Rotation Method: Varimax with Kaiser Normalization 3. Total variance explained: 89.53%

Multivariate Analysis of Variance (MANOVA) was employed to examine whether the experienced eLQA participants perceived the quality framework the same as the prospective participants did in the innovation attributes of trialability/complexity, observability, relative advantage and compatibility. The mean scores of participants’ overall attitudes toward eLQA are shown in Table 8. Participants revealed positive attitudes toward the observability (mean=3.28), relative advantage (mean=3.13) and compatibility (mean=3.19) of eLQA, and stood neutral toward the trialability/complexity (mean=2.97) of eLQA. In other words, the positive observability indicated that the eLQA participants felt that the effects of adopting eLQA could be visible easily, the positive relative advantage also indicated that the adoption of eLQA would bring about better quality in e-learning, and the positive compatibility revealed that the adoption of eLQA was compatible with the development of quality e-learning within participants’ organizations. Finally, the neutral attitude toward the trialability/complexity of eLQA indicated that it was not too difficult or complex for participants to try out the quality evaluation by themselves before submitting an eLQA application. Whether the experience of eLSC and eLCC quality evaluation affects participants’ perception of the quality framework is further examined as follows.
### Table 8. Mean scores of participants’ overall perceptions of eLQA

<table>
<thead>
<tr>
<th>Innovation attributes</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trialability/complexity</td>
<td>2.97</td>
<td>0.72</td>
<td>71</td>
</tr>
<tr>
<td>Observability</td>
<td>3.28</td>
<td>0.89</td>
<td>71</td>
</tr>
<tr>
<td>Relative advantage</td>
<td>3.31</td>
<td>0.80</td>
<td>71</td>
</tr>
<tr>
<td>Compatibility</td>
<td>3.19</td>
<td>0.69</td>
<td>71</td>
</tr>
</tbody>
</table>

### Analysis of eLSC Participants’ Perception

One-way MANOVA was conducted to examine the effects of eLSC experience on participants’ perceptions of the trialability/complexity, observability, relative advantage and compatibility of the quality framework with the significance level of .05. Participants were identified as the experienced eLSC participants group and the prospective eLSC participants group, according to previous experience of eLSC evaluation and experience in managing the quality e-learning service. That is to say, the experienced group not just had the experience in managing e-learning service quality but also the experience of the eLSC quality evaluation. In contrast, the prospective participants were familiar with the eLSC quality framework and had the experience in managing the quality of e-learning service but without the eLSC quality evaluation experience. For achieving the goal of successful diffusion of eLSC to the potential adopters, the experienced participants’ attitudes toward eLSC were hypothesized to be enhanced and reach a higher level of perception than the prospective participants did as measured in the attributes of trialability/complexity, observability, relative advantage and compatibility. In other words, if eLSC was successfully diffused to the adopters, it should bring about positive attitudes in the experienced participants and ensure the continuing use of the quality framework.

The mean scores of the experienced-participants group and the prospective-participants group are shown in Table 9. Both the experienced participants and prospective participants stood a neutral attitude toward the trialability/complexity of eLSC with mean scores of 2.96 and 2.97, respectively. The experienced participants revealed positive attitudes (observability: mean=3.55; compatibility: mean=3.31) toward the attributes of observability and compatibility, and the prospective participants showed neutral-toward-positive attitudes (observability: mean=3.07; compatibility: mean=3.09). Moreover, the prospective participants showed a positive attitude (mean=3.55) toward the relative advantage of eLSC but the experienced participants stood a neutral perception (mean=3.02).

Box's Test of equality of covariance matrices was insignificant (Box’s M = 6.02, $F = 1.688, p = .167$). The homogeneity assumption was sustained. The significant Wilks’ Lambda (Wilks’ Lambda = .584, $p < .001, \eta^2 = .416$) indicated that participants’ eLSC experience affected their perceptions of the quality framework. The main effects were further examined as follows.

### Table 9. Group means of eLSC participants’ perceptions

<table>
<thead>
<tr>
<th>Innovation attributes</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trialability/complexity</td>
<td>Prospective</td>
<td>2.97</td>
<td>.63</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>2.96</td>
<td>.82</td>
<td>32</td>
</tr>
<tr>
<td>Observability</td>
<td>Prospective</td>
<td>3.07</td>
<td>.97</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>3.55</td>
<td>.68</td>
<td>32</td>
</tr>
<tr>
<td>Relative advantage</td>
<td>Prospective</td>
<td>3.55</td>
<td>.55</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>3.02</td>
<td>.96</td>
<td>32</td>
</tr>
<tr>
<td>Compatibility</td>
<td>Prospective</td>
<td>3.09</td>
<td>.65</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>3.31</td>
<td>.74</td>
<td>32</td>
</tr>
</tbody>
</table>

Note. N=71

The MANOVA summary of eLSC experience on participants’ perception of the innovation attributes of eLSC is shown in Table 10. The main effects of the observability and relative advantage were significant (observability: $F_{(1, 69)} = 4.932, p = .030, \eta^2 = .068$; relative advantage: $F_{(1, 69)} = 10.301, p = .002, \eta^2 = .132$). The results indicated that the experienced eLSC participants possessed a higher level perception (mean=3.55) of the observability of eLSC than
the prospective participants did (mean=3.07). In other words, the experienced participants felt that the positive effects of the adoption of eLSC could be observed apparently. In contrast, without the eLSC experience, the prospective participants only stood a neutral attitude toward the observability of eLSC. Therefore, the successful diffusion hypothesis was retained and participants’ positive perception was enhanced through the observability attribute of eLSC.

Furthermore, the prospective participants revealed a higher level perception (mean=3.55) of the relative advantage of adopting eLSC than the experienced participants did (mean=3.02). That is to say, participants without eLSC experience tended to perceive the relative advantage of eLSC positively, but after experiencing the eLSC evaluation, the experienced participants stood a neutral perception. With comparison to the positive attitude possessed by the prospective participants, the experienced participants’ perception of the relative advantage of eLSC could be treated as decreased after experiencing eLSC. Therefore, the successful diffusion hypothesis was rejected and participants’ perception of the relative advantage decreased but sustained at a neutral level after experiencing eLSC.

Finally, the non-significant difference and the mean scores of the trialability/complexity and compatibility indicated that both un-experienced and experienced eLSC participants possessed neutral attitudes toward the trialability/complexity and compatibility of eLSC. In other words, the eLSC experience did not enhance or decrease participants’ perceptions of the trialability/complexity and compatibility. Although the successful diffusion hypotheses were rejected, participants still sustained a neutral attitude toward the attributes of trialability/complexity and compatibility.

In conclusion, although the preparation of supporting evidences for the eLSC quality evaluation usually takes tremendous time and efforts, the enhanced perception of the observability of eLSC indicated that participants were able to see the positive effects of adopting eLSC on the enhancement of e-learning service quality. Therefore, the adoption of eLSC can be facilitated through the observability attribute of eLSC. Moreover, the neutral attitudes toward the trialability/complexity and compatibility suggested that the eLSC quality framework possessed moderate trialability/complexity and compatibility for sustaining eLSC adopters. However, the neutral but decreased relative advantage was suggested to be further enhanced in order to facilitate a successful and pervasive adoption of the eLSC quality framework among e-learning service providers.

### Table 10. MANOVA summary of eLSC experience on participants’ perception

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>eLSC experience</td>
<td>Trialability/complexity</td>
<td>.016</td>
<td>1</td>
<td>.016</td>
<td>.034</td>
<td>.854</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Observability</td>
<td>3.641</td>
<td>1</td>
<td>3.641</td>
<td>4.932*</td>
<td>.030</td>
<td>.068</td>
</tr>
<tr>
<td></td>
<td>Relative advantage</td>
<td>5.668</td>
<td>1</td>
<td>5.668</td>
<td>10.301*</td>
<td>.002</td>
<td>.132</td>
</tr>
<tr>
<td></td>
<td>Compatibility</td>
<td>.937</td>
<td>1</td>
<td>.937</td>
<td>1.965</td>
<td>.166</td>
<td>.028</td>
</tr>
</tbody>
</table>

Note. * p<.05, N=71

### Analysis of eLCC Participants’ Perception

One-way MANOVA was conducted to examine the effects of eLCC experience on participants’ perceptions of the trialability/complexity, observability, relative advantage and compatibility of the quality framework with the significance level of .05. Participants were identified as the experienced eLCC participants and the prospective eLCC participants according to previous experience of eLCC evaluation. The experienced participants were familiar with the eLCC quality evaluation and possessed the experience of eLCC evaluation. Furthermore, the prospective participants were also familiar with the quality framework but without the eLCC quality evaluation experience. For achieving the goal of successful diffusion of eLCC, the experienced participants’ attitudes were hypothesized to reach a higher level of perception as measured in the attributes of trialability/complexity, observability, relative advantage and compatibility than the prospective participants did. Therefore, if eLCC was successfully diffused to the adopters, it should bring forth positive attitudes in the experienced participants and enhance the continuing use of the quality framework.

As shown in Table 11, the mean scores revealed that the prospective participants possessed positive attitudes toward the trialability/complexity (mean=3.20), observability (mean=3.42), relative advantage (mean=3.50) and
compatibility (mean=3.13) of adopting eLCC. Furthermore, the experienced participants also showed positive attitudes toward the observability (mean=3.18), relative advantage (mean=3.17) and compatibility (mean=3.23) of eLCC, and stood a neutral perception (2.79) of the trialability/complexity of eLCC. The experienced participants’ positive attitudes implied that the observability, relative advantage and compatibility of eLCC are adequate for facilitating the adoption of the quality framework. However, the neutral perception of the trialability/complexity suggested that the eLCC quality framework could be further reinforced to increase trialability and reduce the complexity in order to facilitate the adoption of eLCC.

Box's Test of equality of covariance matrices was not significant (Box’s M = 4.64, $F = 1.30$, $p = .272$). The homogeneity assumption was sustained. The significant Wilks’ Lambda (Wilks’ Lambda = .785, $p = .007$ and $\eta^2 = .215$) indicated that eLCC experience affected participants’ perception of the quality framework. The main effects were further examined as follows.

**Table 11. Group means of eLCC participants’ perception**

<table>
<thead>
<tr>
<th>Innovation attributes</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trialability/complexity</td>
<td>Prospective participants</td>
<td>3.20</td>
<td>.47</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Experienced participants</td>
<td>2.79</td>
<td>.80</td>
<td>40</td>
</tr>
<tr>
<td>Observability</td>
<td>Prospective participants</td>
<td>3.42</td>
<td>.90</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Experienced participants</td>
<td>3.18</td>
<td>.87</td>
<td>40</td>
</tr>
<tr>
<td>Relative advantage</td>
<td>Prospective participants</td>
<td>3.50</td>
<td>.98</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Experienced participants</td>
<td>3.17</td>
<td>.61</td>
<td>40</td>
</tr>
<tr>
<td>Compatibility</td>
<td>Prospective participants</td>
<td>3.13</td>
<td>.84</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Experienced participants</td>
<td>3.23</td>
<td>.56</td>
<td>40</td>
</tr>
</tbody>
</table>

Note. N=71

The MANOVA summary of eLCC experience on participants’ perception of the innovation attributes of eLQA is shown in Table 12. The main effects of trialability/complexity and relative advantage were significant (trialability/complexity: $F(1, 69) = 10.809$, $p = .002$, $\eta^2 = .137$; relative advantage: $F(1, 69) = 4.870$, $p = .031$, $\eta^2 = .067$). The significant trialability/complexity main effect suggested that the eLCC participants’ perception of the trialability/complexity attribute of eLCC decreased (from 3.20 to 2.79) and sustained at a neutral level after experiencing the quality evaluation. The decreased perception of the trialability/complexity attribute indicated that the eLCC quality evaluation was not as easy as it expected to try out by the prospective participants. Therefore, the experienced participants only sustained at a neutral level of perception toward the trialability/complexity attribute. Accordingly, participants’ perception of the relative advantage of eLCC also decreased (from 3.50 to 3.17) but sustained at a moderately positive level after experiencing the eLCC quality evaluation. Hence, the successful diffusion hypotheses were rejected for the trialability/complexity and relative advantage attributes, and participants sustained a neutral perception after experiencing the eLCC quality evaluation.

Moreover, the non-significant main effects of observability and compatibility main effects and the mean scores, as shown in Table 11, indicated that the observability attribute of eLCC was moderate and the positive effects of adopting eLCC were easily visible to the e-learning courseware developers, and the eLCC quality framework was compatible with the development process employed and values possessed by the courseware developers. Although the successful diffusion hypotheses were rejected, participants still sustained neutral attitudes toward the observability and compatibility of eLCC after experiencing the quality evaluation.

To sum up, although eLCC participants’ perceptions of the trialability/complexity and relative advantage of eLCC decreased after experiencing the quality evaluation, participants maintained positive perceptions of the observability, relative advantage and compatibility of eLCC and sustained a neutral attitude toward the trialability/complexity eLCC. In other words, participants valued the observability, relative advantage and compatibility attributes of eLCC. Therefore, the positive attributes of observability, relative advantage and compatibility of eLCC would enhance the adoption of the eLCC quality framework. Finally, the neutral perception of the trialability/complexity suggested that the eLCC quality framework could be further reinforced to reduce the complexity and increase trialability in order to facilitate the adoption of eLCC.
Table 12. MANOVA summary of eLCC experience on participants’ perception

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>eLCC experience</td>
<td>Trialability/complexity</td>
<td>4.954</td>
<td>1</td>
<td>4.954</td>
<td>10.809*</td>
<td>.002</td>
<td>.137</td>
</tr>
<tr>
<td></td>
<td>Observability</td>
<td>.639</td>
<td>1</td>
<td>.639</td>
<td>.866</td>
<td>.355</td>
<td>.013</td>
</tr>
<tr>
<td></td>
<td>Relative advantage</td>
<td>2.680</td>
<td>1</td>
<td>2.680</td>
<td>4.870*</td>
<td>.031</td>
<td>.067</td>
</tr>
<tr>
<td></td>
<td>Compatibility</td>
<td>.283</td>
<td>1</td>
<td>.283</td>
<td>.593</td>
<td>.444</td>
<td>.009</td>
</tr>
</tbody>
</table>

Note. * p<.05, N=71

Discussions

Quality assurance is a systematic and comprehensive effort for improving quality. A quality framework can serve as a common platform for organizations, individuals, practitioners and educators. Through the adoption or development of a suitable quality assurance framework, an organization can deal with the quality challenge in a systematic and effective manner. Although there are various quality guidelines, standards, benchmarks and checklists developed by organizations and researchers to respond to the needs caused by the rapid growth of Internet-based learning, the results of the analysis may vary from one instrument to another as well as from one organization to another. Therefore, e-learning quality instruments should be utilized and interpreted with caution (Borotis & Poulymenakou, 2004). Furthermore, the adoption of a quality framework is a long-term process and requires a great deal of organizational change in perception, workflow and culture (Daft, 2006). Changing the mindset of potential adopters or organizations is the first important task in facilitating a smooth process of innovation diffusion. Therefore, it is imperative to examine how the potential adopters as well as the adopters perceive the innovation.

In the present study, the appropriateness of the eLQA quality framework for e-learning was confirmed in terms of validity, reliability and participants’ perceptions. The eLSC quality framework was developed based on the perspective that “quality is how the organization goes about achieving its objectives” and incorporated an in-depth hierarchy constituted with quality aspects, quality criteria and sub-criteria to assess the quality level of e-learning service of an organization from multiple perspectives. The successful implementation of the eLSC quality evaluation not only verified the applicability of the quality framework itself to the e-learning industries, but also aroused general awareness toward quality issues among e-learning providers in the country. As expected, the results of gap analysis indicated that participants perceived their e-learning quality overly for all eLSC quality criteria. This indicated that consultation services such as by-phone or face-to-face were imperative for resolving the perception gaps. Correspondingly, the lowest self-evaluation score and the lowest eLSC evaluation score both fell into the instructional design criterion and indicated that the eLSC participants were aware of their shortage in providing appropriate instructional design to supporting quality e-learning in their products. This shortage may caused by the participants’ inability or unwillingness. For those participants who were incapable of doing so could be facilitated by providing appropriate training on instructional design skills and knowledge. However, the reasons for the unwilling participants might be caused by short development period of e-learning projects, cost-effective consideration, or other factors, and were suggested to be further investigated.

Moreover, the positive observability and compatibility and the moderate trialability/complexity and relative advantage of eLSC confirmed that the eLSC quality framework was appropriately implemented. However, after experiencing the eLSC quality evaluation, the relative advantage of eLSC decreased but sustain at a moderate level. The diminished dominant relative advantage of eLSC could be explained by the overestimate revealed in gap analysis. The frustration could cause an adopter to cease the adoption of eLSC. Therefore, how to sustain the relative advantage of eLSC becomes a top priority issue in improving the eLQA quality framework. Furthermore, the lowest quality level of the sub-criteria determined the overall quality of e-learning in the eLSC quality evaluation. The lack of flexibility in comprising the quality performance of sub-criteria onto a superordinate criterion-level quality level not just caused the criterion-level to become useless in representing the quality level of e-learning but also restrained the applicability of the quality framework. Therefore, it was suggested that the determination of eLSC quality level could be refocused and move back onto the criterion-level in order to represent the quality level of e-learning in fidelity.
Similarly, the successful implementation of the eLCC quality evaluation has aroused general awareness toward quality issues among e-learning courseware developers. There are two major problems evolved, including learner tracking and instructional design, from the analysis. For achieving the adaptive characteristic of e-learning, courseware needs to be design to monitor the learner through out the learning, provide adaptive interaction with the learner accordingly, and update learners’ learning records through the Internet. However, this adaptive feature can only work on an appropriate e-learning platform. Whether the learner tracking gap is caused by technical issue, poor instructional design, cost-effective consideration needs to be further examined. Moreover, the insufficient instructional design quality problem may also be caused by the participants who were incapable or unwilling to do. Focusing on the quality framework issues, the insufficient instructional design may be just caused by participants' misinterpretation of the quality criteria. However, the development of e-learning courseware relies on collaborative works of subject matter experts, teaching experts, instructional designers, educational psychologists, multi-media specialists and web-technology specialists. Meanwhile, all the ingredients need to take effect to suit the educational underpinning in order to bring forth effective learning in the learners. However, the causes of insufficient instructional design quality could be more complicate. Furthermore, the analysis on participants’ attitudes toward eLCC suggested that the eLCC quality evaluation was adequately implemented and possessed positive observability, relative advantage and compatibility, and moderate trialability/complexity. The decreased trialability/complexity and relative advantage of eLCC also indicated the necessity for further improvement. Crystal-clear descriptions, examples and best practices of quality criteria are suggested to make the eLCC quality framework more amiable and easy of access to e-learning developers and organizations.

In addition, this study adopted the construct of innovation attributes to assess participant’s perceptions of the introduced new quality framework for e-learning. Four principal components of innovation attributes were emerged, including trialability/complexity, observability, relative advantage and compatibility, and explained 89.53% of the variance in the study. The results of factor analysis were in accordance with Rogers’ assertion of the interrelated characteristics of innovation attributes (Rogers, 2003) and similar to previous studies (Goldman, 1994; Pankratz, Hallfors & Cho, 2002; Sonnenwald, Maglaughlin, & Whitton, 2001). For those e-learning practitioners, the combined trialability/complexity construct may also support the suggestion of providing sufficient crystal-clear descriptions, examples and best practices of quality criteria to narrow quality gaps and facilitate the adoption of the quality framework. Besides, due to the factors of timing and widespread target population, the present study could only examine the difference between the prospective participants and the experienced participants. It is suggested that further studies could employ a repeated measure design to examine the same participant’s perceptions before and after the quality evaluation in order to get a better fidelity of the change in participants’ perceptions.

There are related issues emerged. First, a quality framework might not suitable for all types of e-learning courseware as well as e-learning organizations. The eLQA quality assurance program needs to be considerate and open any necessary revision in order to meet e-learning participants’ needs. Second, a fair and objective quality evaluation relies on evaluators’ consensus, as well as other participants’ consensus, toward the meanings represented by the quality criteria. Therefore, evaluator trainings and interactions among e-learning participants are essential for emerging and sustaining such consensus. Third, an implemented quality framework might restrict the subsequent development of e-learning. The impact needs to be carefully monitored and studied systematically in order to prevent negative effects that hamper the positive development of e-learning. Fourth, the compatibility of the fixed set of quality standards with the newly developed technological applications, such as the collective intelligence of web 2.0 and the semantic networks of web 3.0, remains an issue. Fifth, a positive perspective of quality assurance indicates pursuing better quality in learning. In contrast, a negative perspective might imply just reaching the minimum requirement. Unfortunately, the final decision-making judgment usually falls on a cost-effective consideration which jeopardizes the development of pedagogically effective courseware in most cases. Finally, the educational effectiveness of the Internet-based learning was not spontaneously ensured through the utilization of the Internet and multimedia technologies. The eLQA quality framework demonstrated adequate reliability and validity, and was capable of assessing learning services and courseware reliably. However, the key measure of quality is “whether a learner can demonstrate the skills, knowledge and competencies set out for them by the organization” (Meyer, 2002). The process-oriented evaluation of eLQA does not measure whether learners’ learning outcomes meet the pre-set objectives of an organization directly. Further studies are suggested to validate the relationships between the process-oriented measures and the outcomes of e-learners.
References


Appendix

Sample Items of the Attitude Questionnaire

Please indicate your level of agreement for each of the following statements according to your perception of the quality framework with 1, 2, 3, 4, and 5 standing for “Strongly Disagree”, “Disagree”, “Neutral”, “Agree”, and “Strongly Agree”, respectively.

Note. The word “eLQA” in the questionnaire items is substitutable by “eLSC” and “eLCC”.

A. Relative advantage
1. With comparison to the previous/current quality assurance method implemented in my organization, eLQA consists of a complete spectrum of quality aspects for ensuring quality e-learning.
   - □ 1, □ 2, □ 3, □ 4, □ 5
2. With comparison to the previous/current quality assurance method implemented in my organization, eLQA provides opportunities to examine the educational meaningfulness for e-learning in depth.
   - □ 1, □ 2, □ 3, □ 4, □ 5
3. With comparison to the previous/current quality assurance method implemented in my organization, eLQA provides effective mechanisms for ensuring the educational meaningfulness of e-learning.
   - □ 1, □ 2, □ 3, □ 4, □ 5

B. Compatibility
1. The quality aspects and criteria of eLQA are consistent with the major factors emphasized by the previous/current quality assurance method implemented in my organization.
   - □ 1, □ 2, □ 3, □ 4, □ 5
2. The quality aspects and criteria of eLQA can be infused in the previous/current quality assurance process implemented in my organization.
   - □ 1, □ 2, □ 3, □ 4, □ 5
3. The process of conducting eLQA is similar to the previous/current quality assurance process implemented in my organization.
   - □ 1, □ 2, □ 3, □ 4, □ 5

C. Complexity
1. The meanings of the quality aspects and criteria of eLQA are stated concisely.
   - □ 1, □ 2, □ 3, □ 4, □ 5
2. The rating method and criteria of eLQA are clearly stated.
   - □ 1, □ 2, □ 3, □ 4, □ 5
3. Following the self-evaluation instruction of eLQA, I can conduct a precise assessment on an e-learning product.
   - □ 1, □ 2, □ 3, □ 4, □ 5

D. Observability
1. The quality improvement derived from the adoption of eLQA is apparent and can be recognized by most project members.
   - □ 1, □ 2, □ 3, □ 4, □ 5
2. The quality levels of eLQA certification are recognized by the public.
   - □ 1, □ 2, □ 3, □ 4, □ 5
3. The eLQA evaluation report provides clarified information for improving e-learning.
   - □ 1, □ 2, □ 3, □ 4, □ 5

E. Trialability
1. eLQA provides sufficient information concerning ways to examine the quality of e-learning.
   - □ 1, □ 2, □ 3, □ 4, □ 5
2. Following the provided information and resources, such as procedure manuals, examples, checklists, and self-evaluation forms, it is easy to try out some parts of or the complete eLQA quality evaluation.
   - □ 1, □ 2, □ 3, □ 4, □ 5
3. Through the try-out of the eLQA quality evaluation, the educational meanings of quality criteria become clear in my mind.
   - □ 1, □ 2, □ 3, □ 4, □ 5
The Model of Strategic e-Learning: Understanding and Evaluating Student e-Learning from Metacognitive Perspectives

Meng-Jung Tsai
Graduate School of Technological and Vocational Education, National Taiwan University of Science and Technology, 43, Sec. 4, Keelung Rd., Taipei, 106, Taiwan // Tel/Fax: (886-2)27376270 // mjtsai99@mail.ntust.edu.tw

ABSTRACT
This paper presents the Model of Strategic e-Learning to explain and evaluate student e-learning from metacognitive perspectives. An in-depth interview, pilot study and main study are employed to construct the model and develop an instrument—the Online Learning Strategies Scale (OLSS). The model framework is constructed and illustrated by four dimensions of characteristics of e-learning environments and three core domains (perceived-skill, affection and self-regulation) of student e-learning strategies. This study also tries to identify and analyze elements corresponding to each domain strategy. Based on this model, the OLSS finally contains 20 items with good construct validity and reliability. This instrument provides a diagnostic instrument for e-learning researchers, system designers, curriculum developers and instructors to evaluate students’ e-learning strategies in their experiment, design and development. This study concludes by presenting several future research directions.

Keywords
E-learning, e-Learning model, Evaluation, Strategic e-learning, e-Learning strategies

Background
During the past decade, researchers, schools and governments around the world have advocated enhancing student learning by using digital tools, i.e. the e-learning. The biggest advantage of e-learning is that it gives students active learning opportunities. Students are believed to be able to gain greater control over their own learning in e-learning compared to traditional learning. Growing numbers of online instructional systems have been developed during the recent years; however, studies of online instructions provided in these systems did not exhibit consistent results in terms of improving or supporting student learning. This may be related to student learning strategies used for Internet-based learning (McCormack & Jones, 1998). When students shift their learning from traditional to online learning environments, they are challenged by different learning and interaction methods. If students adopt effective and efficient approaches for e-learning, they can enhance both their e-learning achievement and their e-learning motivation.

Previous researches have noticed the role of learning strategies in Internet-based learning. It has been observed that student learning strategy is one of the factors impacting student online learning achievement (Shih, Ingevritsen, Pleasants, Flickinger & Brown, 1998). Wallace (2000) and his colleagues indicate that online information seeking is a complex and difficult process for students and developing students’ understanding of content through use of the Internet is a challenge for students and teachers. Tsai and Tsai (2003) further report that student Internet self-efficacy and metacognitive strategies play important roles in student online inquiry learning. Ligorio (2001) considers that the various communication styles integrated into online learning activities are valued only when students are aware of the technologies and tools associated with each communication style. For example, Frank and his colleagues (2003) examined the process of online learning via e-mail for elementary students and concluded that students encountered technological problems and social problems. Technological problems included anxiety regarding using computers for learning, difficulties in using email and the Internet to complete homework, and the difficulty of solving problems when computer systems are down. Regarding social problems, the most significant social problems related to feelings of isolation resulting from online learning. Most elementary students still needed parental help to finish their homework. Lee (2001) examined the styles of learners accustomed to online learning environments and further found that students who recognized online learning may have poor online learning achievement. These literatures imply that online learners are challenged by new problems which they may have never encountered before in traditional learning environments; for example, how to handle the feelings of isolation and how to solve online technological problems by themselves. Recent research explore online inquiry-based learning and claim that higher-level cognitive strategies facilitate student knowledge construction (Salovaara, 2005) and further develop
scaffolding to enhance the development of student metacognitive strategies (Kramarski & Gutman, 2006; Quintana, Zhang & Krajcik, 2005). These researches reveal that new approaches and cognitive strategies may need to be developed particularly for online learning circumstances. Furthermore, positive attitude towards online learning is revealed not sufficient for successful online learning. This suggests that, in addition to affective variables, online learning may be simultaneously influenced by cognitive and behavioral variables. According to the above, students may be required for new learning strategies and skills so that they can become effective and successful online learners.

Features of Internet-based learning environments require online learners for new approaches to achieve their goals or requirements of online learning; however, research literature regarding online learning strategies remains extremely limited. Miller and Miller (2000) pointed out three features of web-based learning environments that differed significantly from traditional learning environments: (1) associative, nonlinear, and hierarchical structure; (2) enhanced multimedia capabilities; and (3) various synchronous and asynchronous communication opportunities. Two problems that limit the use of hypertext environments are unrestricted learner control of sequencing and lack of learner ability to meaningfully integrate unstructured information (Jonassen, 1991). Collis and Meeuwsen (1999) indicated that students need to take responsibility for increased control of learning process and positive attitude to achieve active learning in such an open-ended learning environment as the Internet. They further suggested that learning how to learn on the Internet should include reflection skills, planning skills, study skills, search skills, application skills, and self-evaluation skills. MacGregor (1999) reported that learners adopt different learning approaches in hypertext learning environment. Linn, Davis and Bell (2004) proposed the knowledge integration perspective of e-learning and emphasized on the designs of scaffoldings of idea recognition, connection and monitoring skills. These above technological, cognitive and metacognitive skills are critical for successful e-learning. The inability of students to transfer their learning strategies directly from traditional to online learning environment may result in different performances (Mehlenbacher, Miller, Covington, & Larsen, 2000). Although some researchers (for example, Miller & Miller, 2000) have already noticed the importance of e-learning strategies, an integrated and complete model to profile students’ e-learning strategies is still lack in related literature.

With high degrees of freedom, the Internet-based learning environments may favor students who possess mature computer skills and metacognitive skills (Park & Hannafin, 1993; Tsai & Tsai, 2003). These skills are both explicitly and implicitly exhibited in their learning strategies (Miller & Miller, 2000; Weinstein & Mayer, 1986). Therefore, an integrated and complete model describing student online learning strategies is necessary for deeply understanding students’ e-learning approaches and systematically assessing variables influencing e-learning achievement. Such a model can help researchers explore many research topics including identifying effective learning strategies for online learning; identifying important elements that should be considered as factors when understanding and describing student online learning strategies; identifying the impact of student online learning strategies on their online learning achievements; developing methods for evaluating and diagnosing student mature levels of online learning strategies; and developing methods for helping students acquire new learning strategies. To explore the above research issues, this study proposes a theoretical model for understanding student learning strategies used in online learning environments. In addition, this work presents the development of an instrument for assessing student online learning strategies. Besides, e-learning may involve various digital technologies including the Internet, CD-ROM, PDA, interactive TV or satellite broadcast…and so on. This study discusses e-learning focusing on learning via the Internet involved learning interfaces. Therefore, terms including e-learning, Internet-based learning and online learning are used interchangeably in this study.

Theoretical Framework

Metacognitive Perspectives

Metacognition refers to the self-awareness of individuals about their knowledge and self-understanding, self-control and self-manipulation of the process of their own cognition (Osman & Hannafin, 1992). Students with high metacognitive abilities not only are clearly aware of their learning objectives, but also know effective and efficient approaches to construct knowledge; therefore, such students can monitor their own learning and utilize various learning strategies, thus enhancing leaning achievement and learning motivation (Pressley & Wolshyn, 1995). Park and Hannafin (1993) even observed that students require higher metacognitive skills in less-structured learning environments. Although instructors can accommodate students with different levels of metacognitive skills by
selecting suitable teaching objectives and activities (Miller & Miller, 2000), some researchers (for example, McKeachie, 1988; Weinstein & Mayer, 1986) emphasized on particular curriculum design and instruction to develop student metacognitive abilities and explore the effects of providing scaffolding for learning strategies on learning achievement and motivation. Some strategy-embedded curricula were supported by long term observation and examination (Weinstein, 1994; Weinstein & McCombs, 1998). Therefore, Weinstein (1994; 1998) proposed the concept of “strategic learning” to explain student learning strategies based on metacognitive perspectives.

**Strategic Learning**

The strategic learning (Weinstein, 1994; 1998) is focused on students as active, self-determined individuals who process information and construct knowledge. The model has the learner at its core, and around this core are three interactive components that explain successful learning: **skill**, **will**, and **self-regulation**. **Skill** refers to the various actions or thinking processes related to recognition of key concepts and processes and how meanings are constructed. **Will** indicates individual learning attitude, acceptance of new information, will to concentrate and make efforts, and anxiety toward his/her own learning performance. **Self-regulation** (Zimmerman, 1989) describes how individuals manage their personal learning process, especially how to plan, monitor, focus on and evaluate their own learning. Categorized by the above three components, learning strategies here refers to any **thoughts**, **behaviors**, **beliefs** or **emotions** that facilitate the acquisition, understanding or later transfer of new knowledge and skills. Ertmer and Newby (1996) recognize strategic learning as a feature of expert learning. Strategic learning stresses the need of learners to clearly realize their individual advantages and disadvantages regarding all aspects of strategies to enable them to better manage their learning. Furthermore, to diagnose the strengths and weaknesses of students in relation to the above aspects of learning strategies, a diagnostic assessment instrument, “Learning and Study Strategies Inventory (LASSI), was developed by Weinstein and her colleagues (1988; 1990). Meanwhile, the LASSI is validated in the in-depth analysis conducted by Cano (2006), which involved conceptually grouping LASSI subscales into three categories: **affective strategies**, **goal strategies** and **comprehension monitoring strategies**. Another similar instrument, the Motivated Strategies for Learning Questionnaire (MSLQ), developed by Pintrich and his colleagues (1991) is designed to investigate student motivation and their use of different learning strategies.

Collis and Meeuwsen (1999) have identified key differences among students in their use of the Internet for learning. They began to consider the sorts of skills and approaches to online learning that students need to make creative and constructive use of the learning potential of the WWW. Particularly, they described the challenges of online environments with regard to “learning to learn” components, including articulation and reflection, planning skills, study skills, finding and applying relevant examples, and self-evaluation. For example, the technological aspects of Internet-based learning environments may be unfamiliar to students, preventing students from articulating their experiences. Decomposing an online learning task into a sequence of subtasks and making realistic estimates of the time and effort required for each subtask in online learning environments may be too complex for students to plan and manage their own online learning. New tools and terminology therefore must be mastered to do something as simple as making a margin note. Choosing meaningful information from the Internet and integrating it to learning domains may represent another challenge for all online learners. As a result, a new model may be required to profile the online learning strategies used by students in Internet-based learning environment.

Based on the above rationale, this study proposes the Model of Strategic e-Learning by modifying the construct of strategic learning of Weinstein (1994; 1998) based on the arguments of Collis and Meeuwsen (1999).

**The Model of Strategic e-Learning**

To develop the model, this study begins with answering the following two research questions: (1) What are the distinct characteristics of online learning environment? (2) What are the characteristics of online learning strategies? According to the strategic learning (Weinstein, 1994), the former is analyzed in four dimensions: **nature of e-learning tasks**, **online social supports**, **available online resources** and **online learning systems**. Meanwhile, the latter is analyzed by conducting an in-depth interview emphasizing student online learning strategies in the three domains of **skill**, **will** and **self-regulation**.
Features of Online Learning Environment

Online learning environment is not involved just one environment but several environments including the WWW, e-mail, asynchronous discussion forum (e.g. BBS, mailing list and newsgroup) and synchronous discussion forum (e.g. online chat room, video conference and online games)... and so on. The related literature of Internet-based learning environment (Miller & Miller, 2000; Linn, Davis & Bell, 2004) is analyzed based on the four environmental dimensions which are recognized from the construct of strategic learning (Weinstein, 1994) and stated in the above paragraph. Finally, four characteristics of online learning environment are identified as follows:

1. **Flexible time and space:** The most significant revolution of e-learning is that e-learning overcome the limits of time and place for learning. Online learning tasks can be performed at any time or place provided they have Internet access. Online learning environments thus differ significantly from traditional learning environments, where teachers and students must meet regularly at a specific place and time. Online learners become more free, flexible and convenient in time and location of learning. That is, students gain more locus of control than learning in traditional environments. However, are students prepared for this freedom? Are they aware of their responsibilities of their online learning? Are they able to set goals and make plans for their online learning? Are they able to self-control and self-monitor their learning process via the Internet? Are they able to concentrate on their learning tasks and manage their time for online learning? All of the questions may challenge novice online learners in their control strategies and metacognitive skills such as self-awareness, self-evaluation, self-monitoring, self-control and time management.

2. **Indirect Social Interaction:** The absence of face-to-face interaction is one of most criticized features of learning via the Internet, although it may reduce the anxiety of some students regarding answering questions in traditional classrooms. The isolation of online learning was one of the main frustrations associated with online learning before synchronous communication technologies (e.g. video/audio conference) significantly accelerated the processing of multimedia signals. And asynchronous communication applications (e.g. online discussion board, e-mail, blog, BBS...etc.) are still used popularly in schools. That is, indirect interactions are still the main streams in e-learning. Besides, online social supports may come not only from teachers, peers and friends but also from unknown individuals around the world. The roles of teachers in an online learning environment become more like facilitators and helpers. In particular, the beliefs and expectations of online teachers may not be perceived by students as easily and strongly as in traditional classrooms. The above challenges require students for new strategies to cooperative and negotiate with others via the Internet. They also need to understand the change of teacher’s role and possess new attitudes, motivation and approaches to interact with varied online social supports.

3. **Abundant Information Resources:** The abundance and diversity of information resources is the greatest advantage of Internet-based learning. Students can immediately retrieve information around the world only if they link to the Internet. However, the quality of online information is varies significantly and information stability and authority differs markedly from that in a traditional classroom. Online learners can easily and quickly search for information they need but simply copy and paste it for their assignments. Students need to know not only how to search for information via the Internet but also how to evaluate, integrate and judge the information provided on the Internet. How to process and utilize online information is a new issue for online learners. Therefore, students require mature online information search strategies and mature online information commitment (Wu & Tsai, 2005) for meaningful online learning. In addition, the numerous of online information could result in anxiety towards online information, for example, individuals may feel anxious because they think online information is too much to read. This may be another challenge needed to be handled by some online learners due to the abundant information resource provided on the Internet.

4. **Dynamic Learning Interfaces:** Online learning involves using Internet technology to create a learning platform for interactions among students and teachers. Due to the technology advances rapidly, the interfaces changing with time dynamically is a specific feature of online learning systems. The systems including hardware and software are usually updated or replaced. The interface design and system function significantly influence student motivation, attitude and achievement of online learning. Unstable online learning systems always cause frustration and anxiety with regard to online learning. As a result, it is necessary for students to understand the nature of Internet technology, application of a new system and approaches to handle it. In addition, online learners need to know how to solve problems or what resources (e.g. online assistant) they can ask for help when they encounter frustrations due to learning systems. Finally, positive attitude towards Internet technology is also required for successful online learning.
In summary, online learning environments challenge students with flexible time and space, indirect social interactions, abundant information resources and dynamic learning interfaces. Therefore, different learning strategies should be required for online learners to perform online learning effectively and efficiently.

**Features of e-Learning Strategies**

*In-depth Interview*

To understand online learning strategies adopted by students, an in-depth interview is performed based on the construct of strategic learning. Table 1 summarizes the interview framework.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Potential Elements</th>
<th>Explanation</th>
<th>Interview Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension</td>
<td>Strategies used to process online information and construct meaningful learning</td>
<td>When reading (or comprehending, or searching, or discussing) online teaching materials, what did you do to enhance your achievement and efficiency? Moreover, what did you do when you encountered problems?</td>
<td></td>
</tr>
<tr>
<td>Skill</td>
<td>Internet Skill</td>
<td>Fundamental Internet skills required for online learning</td>
<td>How long have you been using the Internet? Moreover, what are the Internet skills that every e-learner should learn before taking an online course? What skills have you used in this course?</td>
</tr>
<tr>
<td>Self-Awareness</td>
<td>Metacognitive knowledge possessed by individuals regarding themselves as online learners and regarding the nature of e-learning</td>
<td>In your opinion, is there any difference between e-learning and traditional learning? If yes, what are the differences? Moreover, what roles do students play in both curricula?</td>
<td></td>
</tr>
<tr>
<td>Will</td>
<td>Attitude</td>
<td>Attitudes towards the Internet and online learning</td>
<td>What do you like and dislike in such an online course? Why?</td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>Motivation of online learning</td>
<td>Would you take such a course again in the future if given the opportunity? Why or why not?</td>
</tr>
<tr>
<td>Self-Regulation</td>
<td>Self-monitoring strategies adopted to deal with the requirements of online courses</td>
<td>What approaches did you use to monitor, manage or regulate your online learning? For example, did you make plans for online learning? How did you control the plans you made?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Concentration</td>
<td>Self-control strategies used to concentrate on e-learning</td>
<td>What approaches did you adopt to allow yourself to focus on your online learning?</td>
</tr>
</tbody>
</table>

In order to establish a strategic-learning interview framework, some potential elements are identified and initially categorized into the three domain strategies: skill, will and self-regulation. For example, the skill domain strategies involve the comprehension skills, Internet skills and self-awareness of online learners. Table 1 lists explanations and interview questions corresponding to each element.

Fourteen university senior students serve as subjects for the in-depth interview. The subjects were enrolled in a one-semester-long online course called “Computers and Education” conducted at a Taiwanese university. In this course, students were required to learn via the Internet individually and collaboratively with the exception of two face-to-face classroom meetings conducted for purposes of course introduction and evaluation. At the end of the semester, the subjects were interviewed according to the guidelines and questions listed in Table 1. All of the interviews were recorded and the transcripts were analyzed to examine or identify the elements of each domain strategies or even to identify specific e-learning domains. The concrete descriptions of the subject e-learning strategies were drawn out.
and grouped under a related element and domain. The study data were also used as a key resource of the item pool of the instrument developed and described later in this study.

**Elements of Online Learning Strategies**

Based on the results of the in-depth interviews, three domain strategies specifically related to e-learning were identified to produce the perceived-skill, affection and self-regulation domains. All domains and elements with corresponding descriptions and examples of students’ responses in the in-depth interview are summarized in Table 2.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Elements</th>
<th>Descriptions</th>
<th>Examples of Interview Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension</td>
<td>Constructed meaningful learning in an online learning environment is more complex than in a traditional learning environment. Specifically, achieving this requires high levels of cognitive abilities to process the information provided in an open-ended learning system such as the Internet. Many metacognitive skills are implicit and difficult to detect for online learners.</td>
<td>Sometimes I just don’t understand what is displayed on screens in an online course. I try to catch the key points by taking notes for online learning. I watched the online course videos again and again. I look through titles or links before clicking any link for details.</td>
<td></td>
</tr>
<tr>
<td>Internet Skill</td>
<td>Internet skill</td>
<td>Fundamental Internet skills required for online learning. This includes online search skills, online discussion skills, Internet file transfer skills, and so on. These skills are basic requirements for successfully undertaking online learning and the learners can easily be aware of these skills.</td>
<td>Online learners need to know how to control the system. One should learn how to use computers and Internet before online learning. When I cannot talk clearly in texts on MSN, I use Internet phone (e.g. skype) for online discussion.</td>
</tr>
<tr>
<td>Self-Awareness</td>
<td>Metacognitive knowledge about self as online learners and about the nature of e-learning may be related to online learning experiences. Such knowledge is difficult for novice online learners to figure out by themselves particularly knowledge regarding their obligations and responsibilities. And this is rarely happened in a traditional learning environment.</td>
<td>Online learning is active learning. We should spend more time on searching for information. Online learning is free in location. We don’t have to come to school campus. It is convenient for students who live far from schools. I am not clear about the teacher’s requirements for this online course.</td>
<td></td>
</tr>
<tr>
<td>Attitude</td>
<td>Student attitudes towards the Internet and online learning are intimately related. Student perceptions of the Internet play a role in their willingness to use it for learning. Learning attitude in traditional learning environments is not such heavily dependent on a particular media as the Internet.</td>
<td>I think online learning is fancy. I don’t like online learning, because I cannot ask teachers for questions. I don’t like online learning because I don’t like type. I think online learning can help me learn more than traditional learning.</td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>Motivation for online learning may be more varied and complex than that for traditional learning. Moreover, locus of control in online learning environments</td>
<td>Since I am curious about online learning, I focus well on the course. I take this online course because I</td>
<td></td>
</tr>
</tbody>
</table>
Anxiety of online learning is also related to student Internet using experience and online learning experience. The updating or changing characteristics of e-learning systems may create pressure in relation to online learning.

I worry about my computer skills for online learning.
Online learning brings me pressure.
I am afraid of online learning.
I worry about how to communicate with others in online learning.

Monitoring the process of an online learning task requires a full understanding of associated subtasks and related resources so that learners can set proper goals and self-monitor their progress towards those goals. Students holding lower self-monitoring skills are more likely to have online disorientation problems. This is also a challenge for novice learners.

I always notice the latest news of the online course.
I link to the discussion board everyday in order to update the progress of our group discussion.
I speed up my online studying when I think it is almost delayed in progress.

Online learning is more self-directed than traditional learning. Without many social scaffoldings or supports from teachers and peers, students require active time management skills for effective and efficient online learning.

I follow teachers’ course schedule for online learning.
I turn in my online homework in time.
I randomly link to the online course website without any schedule.

The abundance of multimedia interactions provided in Internet environments distracts online learners from their learning more than in traditional environments. The cognitive load on novice online learners thus is increased. Online learners are required to be acquainted with skills to prevent themselves from numerous attractions of online multimedia or entertainment.

Online games and MSN distract me from online learning.
Once I login to an online course website, I will not logout it before I finish a learning unit.
When I login the online course, I open one window at a time.
I often login to the online website in a quiet environment.

First, the skill domain is modified to become the perceived-skill domain because limited concrete descriptions or examples regarding comprehension and self-awareness can be drawn from their transcript data. This difficulty of detecting the implicit metacognitive skills of individuals may be more significant in e-learning environments due to the cognitive overloading to process online information. Second, since anxiety toward e-learning is strongly displayed in novice Internet users and novice online learners, an anxiety element has been included in the will domain. Therefore, the will domain has been replaced by the affection domain for a broader definition. Finally, time management element is drawn from self-monitoring in the self-regulation domain. This study suggests that these three identified domains and their relevant elements can be actively utilized by students to interact with the e-learning environment. Therefore, the three revised domains and corresponding elements have been all considered as key components to construct the model.

**The Model**

Based on the above analysis of e-learning environments and the in-depth interviews regarding student e-learning strategies, the Model of Strategic e-Learning is conceptualized in Figure 1. First of all, the definition of e-learning in this model indicates the activities and processes through individuals acquire knowledge, skill and attitudes by using...
various digital tools. It is also referred to as Internet-based learning or online learning in this study. Under this broad definition, examples of e-learning include discussing homework with peers via email, searching for information on the Internet, reading online articles, participating in online discussions, and taking distance online courses.

Secondly, the rectangle (in Figure 1) represents an e-learning environment and its four sides represent its four features: **flexible time and space**, **indirect social interactions**, **abundant information resources** and **dynamic learning interfaces**. There are four nodes within the rectangle in this model. The central node represents individual e-learners. Meanwhile, the three surrounding nodes represent the three domains of e-learning strategies which have been identified in the above student in-depth interviews. Thirdly, since individual e-learners can actively control these domain strategies to interact with the e-learning environment, there are three edges linking the central node (meaning the learner) with the three other nodes (meaning **perceived-skill**, **affection** and **self-regulation** domain strategies). The corresponding elements within the three domain strategies are listed aside the three nodes. Finally, bi-directional arrows exist between each pair of nodes, indicating the existence of interactions among the three domain strategies. For example, an individual with higher Internet self-efficacy may have better online search strategies, and a better time management strategy may be related to a better attitude toward e-learning.

In brief, this model can be used to profile how students interact with the complex e-learning environments when they are involved with Internet-based learning. In order to further investigate and modify this model in the future, this study developed an instrument for examining student e-learning strategies. The details are presented in the following section.

**Online Learning Strategies Scale (OLSS)**

Based on the above model, this study developed the Online Learning Strategies Scale (OLSS) as a research instrument to examine student e-learning strategies and provide feedback for modifying the model. The proposed instrument was developed in two stages: a pilot study and a main study. The pilot study involved the collection of a pool of candidate items for this instrument and obtaining preliminary testing statistics. The main study further modified the scale based on the results of a pilot study and further investigates its validity and reliability.
Pilot Study

In order to create a pool of candidate items of OLSS, content analysis was conducted to analyze the transcript data of the in-depth interviews. All clear descriptions and concrete statements regarding student behavioral skills, cognitive skills, metacognitive skills and knowledge, and affective responses reflected or used in online courses in relation to various interview questions are selected as candidate items. Finally, 36 statements were identified as candidate items for the OLSS instrument. Each statement was designed to be measured with a 5-point Likert scale ranging from 1 (not like me at all) to 5 (very much like me).

A survey (Tsai, 2007) involving the 36 items was administered to 136 senior high school students who participated in a six-week Internet-based learning activity as part of their required “Earth Science” course. A total of 118 valid questionnaires were collected. Factor analysis was used to examine structural validity. Furthermore, reliability analysis was used to assess internal consistency. Finally, a total of 22 items were included for further modification and development, and were grouped into the following seven subscales: anxiety, attitude, time management, study aides, self-awareness, Internet literacy, and concentration. Although these items explained 73.15% of the total variance, their reliabilities ranged from 0.53 to 0.85 for subscales. Table 3 illustrates the item numbers and subscale sample items for the initial version of the OLSS.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Reliability</th>
<th>Items</th>
<th>Sample Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anxiety</td>
<td>0.84</td>
<td>4</td>
<td>Online learning makes me very nervous.</td>
</tr>
<tr>
<td>2. Attitude</td>
<td>0.85</td>
<td>4</td>
<td>I think online learning is now important for students.</td>
</tr>
<tr>
<td>3. Time Management</td>
<td>0.80</td>
<td>3</td>
<td>I finish my online homework before the deadline.</td>
</tr>
<tr>
<td>4. Study Aides</td>
<td>0.75</td>
<td>3</td>
<td>I take online notes using word processors or graphic tools.</td>
</tr>
<tr>
<td>5. Self-Awareness</td>
<td>0.71</td>
<td>4</td>
<td>I understand the responsibilities of online learners.</td>
</tr>
<tr>
<td>6. Internet Literacy</td>
<td>0.78</td>
<td>2</td>
<td>I can search for information via the Internet.</td>
</tr>
<tr>
<td>7. Concentration</td>
<td>0.53</td>
<td>2</td>
<td>I am easily deterred from online learning by music or MSN.</td>
</tr>
<tr>
<td>Total</td>
<td>0.82</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

Several suggestions (Tsai, 2007) were made based on the pilot survey results as follows: (1) modify or replace items #32 (I am always late in submitting my online homework) and #27 (I am easily distracted from online learning by music or MSN) to improve the low reliability (alpha=0.53) of the concentration subscale; (2) replace and move item #7 (I am able to participate in online discussions) from the self-awareness subscale to the Internet literacy subscale for further clarification; (3) include more items regarding intrinsic motivation and self-monitoring in the pool; and (4) improve the balance of the item numbers across subscales. Finally, it is necessary to reexamine the validity and reliabilities after modifying the initial OLSS.

Main Study

Based on the above suggestions for further study, the 22 items extracted from the pilot study are modified as follows. Item #32 has been replaced and two new items have been added to the concentration subscale. Furthermore, item #7 has been moved from self-awareness subscale to Internet self-efficacy subscale (renamed the Internet literacy subscale) and a new item (I am not sure about the teacher’s expectation of my online learning) is added to the self-awareness subscale. Motivation and self-monitoring subscales, each comprising four items, are added to the scale. Minor revisions are made to improve clarity and some new items are added to the study aides and Internet self-efficacy subscales to maintain balance across subscales. A revised version including 36 items classified under nine aspects thus is obtained and used for further examination. Again OR Additionally OR Furthermore, each item is measured using a 5-Likert scale ranging from 1(not at all like me) to 5(very much like me).

To validate the revised OLSS, another survey based on the revised 36 items was administered to 400 college students who had online learning experiences during the past three years and volunteered to answer the online survey. Because the survey was conducted on the Internet, extremely rigorous examination was used to discriminate invalid questionnaires. Therefore, only 261 questionnaires were left and used for data analyses in this investigation. A Principal Component factor analysis with Varimax rotation was used to examine the structural validity of the survey.
And a reliability analysis of Cronbach alpha was employed to assess its internal consistency. Table 4 lists the results of the rotated factor loadings and reliability analyses.

Table 4. Rotated factor loadings and Cronbach’s α values for OLSS (20 items)

<table>
<thead>
<tr>
<th>Item No</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subscale 1: Motivation (α=.86)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.788</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>.744</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>.740</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>.705</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.657</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 2: Self-Monitoring (α=.67)</strong></td>
<td></td>
<td>.774</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>.643</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.629</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.535</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 3: Internet Literacy (α=.67)</strong></td>
<td></td>
<td></td>
<td>.787</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>.703</td>
<td>.625</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td>.535</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 4: Internet Anxiety (α=.70)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.798</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td>.749</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td>.696</td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 5: Concentration (α=.70)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.794</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.668</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.591</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.528</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Eigen Value</strong></td>
<td>3.468</td>
<td>2.453</td>
<td>2.123</td>
<td>2.110</td>
<td>2.006</td>
</tr>
<tr>
<td><strong>% of Variance</strong></td>
<td>17.340</td>
<td>12.264</td>
<td>10.613</td>
<td>10.550</td>
<td>10.032</td>
</tr>
<tr>
<td><strong>Total Reliability α= 0.86; Total variance explained is 60.80 % (N=261)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 clearly reveals that five factors are extracted from the factor analysis, and these factors are then termed motivation, self-monitoring, Internet literacy, Internet anxiety and concentration subscales, including 5, 4, 4, 3, and 4 items, respectively. The final version of the OLSS thus contains a total of 20 items with a total explained variance of 60.80%. The reliability alpha is 0.86 for the total scale and ranges from .67 to .86 for all subscales. This means that the final OLSS has better internal consistency than the initial version of OLSS (0.82 for the total and from .53 to .85 for the subscales). Additionally, the numbers of items are more balanced across the subscales in the final version of the OLSS. Although the total explained variance (60.80%) is lower than for the initial version (73.15%), it still exceeds the level of acceptability (> 60%). All 20 items of the final version of OLSS are shown in Table 5. Each item is measured by a 5-Likert scale ranging from 1 (not at all like me) to 5 (very much like me), and thus the total score should range from 20 to 100. Items 6, 9, 12, 26 and 28 should be scored in reverse.

To summarize, the final version of the OLSS successfully modifies the initial OLSS. It not only draws out the motivation and self-monitoring components which are important but do not appear in the initial version but also increases the reliability of the concentration subscale from 0.53 to 0.70. In addition, the final version further clarifies the Internet literacy subscale and improves the balance of the item numbers across subscales. Although some subscales in the initial version do not appear in the final version, some of the items belonging to those subscales are categorized meaningfully in another related subscale. For example, item #10 is still included in the final version but shifts from the time management subscale to the concentration subscale. This phenomenon may occur because the
items of time management subscale in the initial OLSS may not have sufficient loading for an independent factor in
the final OLSS. Item #10 holding stronger relationship with the concentration subscale than with the initial time
management subscale is thus combined into the concentration subscale simultaneously with the time management
subscale disappearing from the final version. Actually, both time management and concentration belong to the self-
regulation domain in the model; therefore, it is not surprising to find that item #10 shifts from the time management
to the concentration subscale. Besides the time management subscale, the attitude, study aids and self-awareness
subscales in the initial version all display similar situations in the final OLSS. Therefore, the reduced number of
subscales in the final version is reasonable and acceptable.

<table>
<thead>
<tr>
<th>Item no*</th>
<th>Subscale</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Motivation</td>
<td>I am attracted to online learning</td>
</tr>
<tr>
<td>22</td>
<td>Motivation</td>
<td>I am excited about learning on the Internet</td>
</tr>
<tr>
<td>29</td>
<td>Motivation</td>
<td>I like online learning</td>
</tr>
<tr>
<td>35</td>
<td>Motivation</td>
<td>I think online learning is convenient</td>
</tr>
<tr>
<td>7</td>
<td>Motivation</td>
<td>I am curious about online learning</td>
</tr>
<tr>
<td>17</td>
<td>Self-Monitoring</td>
<td>I make study plans for online learning</td>
</tr>
<tr>
<td>18</td>
<td>Self-Monitoring</td>
<td>I am able to use online learning aids</td>
</tr>
<tr>
<td>14</td>
<td>Self-Monitoring</td>
<td>I change my online learning plans as necessary</td>
</tr>
<tr>
<td>3</td>
<td>Self-Monitoring</td>
<td>I am able to take online notes using word processors or graphics tools</td>
</tr>
<tr>
<td>5</td>
<td>Internet Literary</td>
<td>I am able to send and receive emails</td>
</tr>
<tr>
<td>20</td>
<td>Internet Literary</td>
<td>I am able to search for information on the Internet</td>
</tr>
<tr>
<td>11</td>
<td>Internet Literary</td>
<td>I understand the differences between online learning and traditional learning</td>
</tr>
<tr>
<td>32</td>
<td>Internet Literary</td>
<td>I am able to download files from the Internet</td>
</tr>
<tr>
<td>28*</td>
<td>Internet Anxiety</td>
<td>Online learning system always make me feel frustrated</td>
</tr>
<tr>
<td>27*</td>
<td>Internet Anxiety</td>
<td>I do not have a good understanding of my progress in online learning</td>
</tr>
<tr>
<td>9*</td>
<td>Internet Anxiety</td>
<td>I am worried about my achievements in online learning</td>
</tr>
<tr>
<td>12*</td>
<td>Concentration</td>
<td>I am easily distracted from online learning by music or MSN</td>
</tr>
<tr>
<td>21</td>
<td>Concentration</td>
<td>I avoid anything that could interrupt my online learning</td>
</tr>
<tr>
<td>6*</td>
<td>Concentration</td>
<td>I cannot focus on my online learning</td>
</tr>
<tr>
<td>10</td>
<td>Concentration</td>
<td>I make good use of my time for online learning</td>
</tr>
</tbody>
</table>

Note. * The item number indicates the item order in the second pool of candidate items (a total of 36 items).
* Reversed scoring items: 6, 9, 12, 27, 28

**Students’ OLSS Scores**

The main study further investigated students’ OLSS scores and their background information. Table 6 summarized
the mean scores, item numbers and average scores for each item in subscales of OLSS. Each average score is
calculated by dividing the mean by the item number for each subscale. The average scores range from 3.17 to 4.36 in
a 5-likert rating scale. This indicates that in average these online learners hold above medium levels of online
learning strategies. The above medium scores are reasonable since the subjects of the main survey were selected
from those who had online learning experience. In addition, the score of Internet Literacy (4.36) is significantly
higher than the total average score (3.66) and the score of Concentration (3.17) is quite lower than the total average
score. This suggests that the online learners possess good Internet literacy skills but only have fair concentration
strategies for online learning. And this result is not surprising when further examining the subjects’ weekly online
hours for general purposes and weekly online hours for learning purpose. The former is 31.43 hours and the later is
only 3.2 hours in average. Furthermore, in average, the 261 subjects were 23.08 years old with 7.96 years of Internet
using experience and had taken 1.4 online courses.

In addition, this study also explored if there was any significant difference in OLSS scores between male and female
students. Table 7 summarized results of the ANOVAs of the OLSS scores between genders. There were 130 males
and 131 females in the 261 online learners. The results showed that there was not any significant difference between
male and female students’ online learning strategies in all dimensions. This means that both male and female online
learners had about the same levels of online learning strategies. However, one thing should be noticed again is that
these subject were not randomly selected for this study, therefore, further investigations regarding gender differences in online learning strategies are still needed in order to draw a conclusion.

<table>
<thead>
<tr>
<th>Table 6: The average item score of OLSS subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLSS</strong></td>
</tr>
<tr>
<td>Motivation</td>
</tr>
<tr>
<td>Self-Monitoring</td>
</tr>
<tr>
<td>Internet Literacy</td>
</tr>
<tr>
<td>Internet Anxiety</td>
</tr>
<tr>
<td>Concentration</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Note. N=261

<table>
<thead>
<tr>
<th>Table 7: The ANOVAs of OLSS scores between male and female students</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLSS</strong></td>
</tr>
<tr>
<td>Motivation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Self-Monitoring</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Internet Literacy</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Internet Anxiety</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Concentration</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Conclusions

This study proposed the Model of Strategic e-Learning to explain student online learning from metacognitive perspectives. The model identifies four characteristics of online learning environments, namely: *flexible time and space, indirect social interactions, abundant information resources and dynamic learning interfaces*. Three domains of e-learning strategies are identified in the mode, namely the *perceived-skill, affection and self-regulation* domains. Potential elements and corresponding features in each domain strategy are also illustrated and elaborated in this work. *Self-regulation* domain is especially important and appropriate for online learners (McManus, 2000), because they need considerable enormous control over their time schedule, and how they approach online learning.

Based on the above model, the Online Learning Strategies Scale (OLSS) is developed as an instrument to evaluate student online learning strategies. The final version of the OLSS is composed of 20 items categorized under five subscales as follows: *motivation, self-monitoring, Internet literacy, Internet anxiety and concentration* subscales, all of which have good reliability. The *motivation* and *Internet anxiety* subscales investigate student *affection* domain strategies, while the *Internet literacy* subscale assesses *perceived-skill* domain strategies, and the *self-monitoring* and *concentration* subscales examine *self-regulation* domain strategies. Some may argue that self-reporting scale is inappropriate for measuring individual metacognition, especially for skills and self-regulation; however, it is possible to identify various cognitive and motivational profiles related to student behavior (Pintrich et al., 1991).

Discussions

The Model of Strategic e-Learning provides a framework for systematically analyzing student online learning strategies. The interfaces of online learning systems are rapidly changing and continuously developing. How online learners handle their own learning and the approaches they use to interact with specific learning systems are
important issues in evaluating the impact of specific e-learning systems on student online learning. The proposed model provides a general scope for understanding and analyzing student online learning and learning strategies. Furthermore, the Online Learning Strategies Scale (OLSS) can be used to examine student online learning strategies. OLSS provides e-learning researchers, system designers, curriculum developers, instructors and even learners themselves with a diagnostic instrument for understanding the advantages and disadvantages of online learning for students, as well as e-learning researchers, system designers and online curriculum developers. OLSS scores can provide feedbacks for e-learning system designers, curriculum developers and instructors before or after they design or develop systems or curriculum. These scores can even serve as feedback that can help online students to obtain greater self-awareness of their own online learning.

Based on the Model of Strategic e-Learning, the study developed the OLSS for evaluating students’ e-learning strategies. However, in this current stage, the OLSS may not fully describe the proposed model and the model may need to be further modified based on repeated surveys. For example, the domain of perceived-skill strategies, especially the comprehension strategies was omitted largely from the final version of OLSS. This may due to the comprehend strategies, not like the Internet skills, are too implicit to be self-detected by learners themselves. Or it could suggest that the comprehension strategies dealing with comprehension monitoring may be closer to self-regulation domain than perceived-skill domain. Therefore, further examinations by OLSS and modifications of the model are needed in the future.

Future studies should also explore the relationships between student online learning strategies and their online learning achievements. In addition, it is important to examine the role played by individual differences in student online learning strategies, for example, the influences of the epistemological views of students, or the relationships of those views with online learning strategies. Finally, the learning strategies included in OLSS are general online learning strategies, and effective learning strategies may be discipline specific. Therefore, further research is required to investigate the appropriateness of the proposed instrument for various online learning activities such as online searching and online discussion.

Acknowledgments

This paper is financially supported by the National Science Council of Taiwan under the projects of NSC94-2524-S-305-001 and NSC95-2511-S-305-001-MY3. Special thanks go to the journal reviewers for their valuable comments and suggestions.

References


48
An Automatic Multimedia Content Summarization System for Video Recommendation

Jie Chi Yang*, Yi Ting Huang*, Chi Cheng Tsai*, Ching I Chung* and Yu Chieh Wu**

*Graduate Institute of Network Learning Technology,
**Department of Computer Science and Information Engineering,
National Central University, No.300, Jungda Rd., Jhongli City 320, Taiwan // Tel: +886 3 4227151 ext.35414, Fax: +886 3 4275336 // *{yang, coral, ken, cecilia}@cl.ncu.edu.tw, **bcbb@db.csie.ncu.edu.tw

ABSTRACT

In recent years, using video as a learning resource has received a lot of attention and has been successfully applied to many learning activities. In comparison with text-based learning, video learning integrates more multimedia resources, which usually motivate learners more than texts. However, one of the major limitations of video learning is that both instructors and learners must select suitable videos from huge volume of data. The situation is worsened by the fact that video content is usually displayed linearly. In order to solve this problem, we propose a multimedia content summarization and adaptable recommendation framework which is able to extract summaries from raw videos and recommend them to learners automatically. By following the characteristics of multimedia, the generated summary contains both important abstracts and corresponding images, and can be accessed online. Only suitable videos are selected for recommendation based on user profiles. The proposed system is evaluated and compared to text-based learning in terms of ARCS model. The results demonstrate that the proposed video summarization and recommendation framework was not only positive with regard to motivating learners, but also enhanced the video learning experience significantly. Positive results are also found in relation to system usage and satisfaction.

Keywords
Video summarization, Video recommendation, Learning material recommendation, Multimedia, Video learning, ARCS model

Introduction

Multimedia-based learning has recently become a promising instructional resource. According to dual-coding theory, the human brain deals with imagery representation better than verbal representation (Clark & Paivio, 1991; Paivio, 1971). Many studies report that multimedia content is often more useful for learning and teaching than traditional text-based learning (Mackey & Ho, 2008; Mayer & Moreno, 2002; Rose, 2003). There are many types of multimedia learning materials, but video is the most representative and popular one. This is because video integrates many multimedia resources, such as text, images, sound, and speech. Based on the theory of constructivism, video provides a context wherein learners can construct their own knowledge (Brown, Collins, & Duguid, 1989). Several studies demonstrate that video is a suitable material for context-based learning (Choi & Johnson, 2005; Choi & Johnson, 2007). In addition, video learning is an effective way of providing motivation, keeping attention, and giving satisfaction to the learner (Choi & Johnson, 2005; Choi & Johnson, 2007; Mackey & Ho, 2008). However, there are some limitations that exist in video learning. For example, choosing a suitable video for instructors and learners from the rapidly growing number of videos can be a problem. A similar problem arises when recommending videos to learners and instructors. Moreover, when learners get a great deal of information in a short period of time, it causes cognitive overload (Pass & van Merrienboer, 1994; Sweller & Chandler, 1994; Sweller, van Merrienboer & Pass, 1998). The usefulness of video learning is obvious, but it is necessary to enhance the recommendation mechanism for learners in order to facilitate multimedia learning.

Unlike texts, watching videos requires much more time since video content is usually displayed linearly. Summarization is used to preserve the most informative parts of the source content. Therefore, video summarization is essential for enabling the learner to skim through video content. With the rapid growth of the video industry, acquiring the appropriate video from a huge database is a difficult task. Adaptive video recommendation could be a way to deal with this situation, as this system is specifically designed to help learners filter information. In essence, the combination of summarization and recommendation is helpful in reducing cognitive overload with regard to video learning. Traditionally, these tasks (i.e., content summarization and recommendation in e-learning) are done...
manually, which is a very demanding and time-consuming process. Consequently, there is a strong demand for an automatic video summarization and recommendation system.

Recently, recommendation systems have been applied to some products and information databases by making adaptive suggestions based on previous examples of a user’s preference (Melville, Mooney & Nagarajan, 2002; Mooney & Roy, 2000; Wang, Tsai, Lee & Chiu, 2007). A recommendation system for learning material can provide objects easily and efficiently thereby enhancing learning activities. Without recommendation mechanisms, learners would spend more time selecting suitable learning objects and less time involved in the actual activity of learning. Several studies have shown (Tsai, Chiu, Lee & Wang, 2006; Wang et al., 2007) that automatic recommendation mechanisms that refer to learner profiles can promote the accuracy of learning object recommendation. Nevertheless, these recommendation mechanisms are only suitable for structured or semi-structured data (Popescul, Ungar, Pennock & Lawrence, 2001; Tsai et al., 2006; Wang et al., 2007). In other words, these systems may not work well with raw videos and raw texts. On the other hand, recommendation systems are also applied to movie or TV recommendations (Alspектор, Kolcz & Karunaithi, 1998; Basu, Hirsh & Cohen, 1998; Cotter & Smith, 2000; Melville et al., 2002). However, these studies tend to apply very limited sets of features such as the movie title, the director, keywords, and actors, as well as like-minded user ratings. The feature that is lacking in all of these studies is the inclusion of speech content, which contains a substantial amount of information relating to the video itself. In other words, the studies mentioned above ignore important content within learning materials.

Automatic summarization is an important research topic, especially in relation to automatic text-based and video-based summarization. Text-based summarization research, such as the Document Understanding Conference (DUC) (http://duc.nist.gov/), aims at extracting important sentences from source documents. These techniques focus on generating summaries from news-like articles (i.e., newspaper and newswire data) (Dang, 2006; Dang, 2007) which are usually shorter and more coherent than video stories. Moreover, a video story usually contains multiple sub-topics. On the contrary, video-based summarization research, such as the TRECVid workshop (http://www-nlpir.nist.gov/projects/t01v/), aims at extracting key-frames and shots from source videos (Over, Ianeva, Kraaj & Smeaton, 2005; Over, Ianeva, Kraaj & Smeaton, 2006), offering a sketch that contains a description of an object (such as color, shape, or motion) (Liu & Li, 2002; Milrad, Rossmannith & Scholz, 2005; Over et al., 2005; Over et al., 2006). This technique is often used in surveillance systems (Osadchy & Keren, 2004; Piriou, Bouthemy & Yao, 2006) and medical videos (Fasquel, Agnus, Moreau, Soler & Marescaux, 2006). Nevertheless, these types of summarization may be not useful for learners due to the neglect of video content. Furthermore, the traditional video-based summarization is not generally used for educational purposes.

In order to help learners choose videos that are suitable for specific learning activities, two issues must be addressed: (a) the appropriate summarization technique must be able to extract semantic information from video content; and (b) the appropriate system must be able to recommend a suitable video to learners from a huge database. The research regarding both of the above points is very limited, and few attempts have been made to apply summarized information to recommendation mechanisms. For example, MovieLens is a well-known movie recommendation website (http://movielens.umn.edu). It generates personalized recommendations on the basis of a user preference. Nevertheless, the recommendation information provided by the site lacks an integral plot summary. Therefore, users might not be able to browse a sufficient amount of information in order to determine whether the movie is related to the knowledge that interests them. YouTube (http://www.youtube.com/), while employing a different system, still falls short with regard to providing videos that are appropriate for specific learning activities. This famous online video streaming service allows anyone to view and share videos that have been uploaded by others. Users can get videos by searching keywords on the website. Unfortunately, users will likely spend a lot of time looking for related videos through the search mechanism rather than receiving relevant information from the recommendation mechanism. Due to such situations, attaining knowledge efficiently through videos may prove to be a difficult task that could even lead to a decrease in motivation on the part of the learner.

Motivation is an important factor for learning. The ARCS model of motivation was formed in response to the necessity of finding more useful ways of understanding the major factors relating to the motivation to learn (Keller, 1983; Keller & Kopp, 1987). This model identifies four major factors: attention, relevance, confidence and satisfaction. All of these factors must be fulfilled if a learner is to become and remain motivated (Dick, Carey & Carey, 2001). Based on the effectiveness of multimedia learning, we hope to develop a video recommendation system that attracts the learner’s attention, recommends relevant videos, and effectively promotes learner confidence and satisfaction.
It is clear that multimedia learning is useful for learners, but there is not a customized tool or mechanism for multimedia learning that perpetuates learner motivation. In this paper, we extend our previous works (Huang, Tsai, Chung, Shen, Yang & Wu, 2007; Tsai, Chung, Huang, Shen, Wu & Yang, 2007) and present an automatic multimedia content summarization and adaptable recommendation system, called Video Content Summarization for Recommendation (VCSR), that auto-recommends suitable multimedia material with the aim of encouraging learners to watch and assimilate knowledge within the framework. The proposed system first extracts video content as a summary and collects corresponding frames from the source. These materials are combined into a hypermedia document and auto-recommended to learners. The system also sends the hypermedia document as email (multimedia-based email) to learners in response to their profiles. Unlike traditional recommendation methods, the system not only recommends video titles, but also includes important extracted content that contains a video summary and corresponding video clips. The system can extract information rapidly from a large database of videos, saving time for the user. Moreover, the system can recommend video material to learners related to what they wish to study. Thus, learners can quickly use the new video information acquired instead of receiving a lot of unnecessary information.

In order for these objectives to be achieved, this paper is structured as follows: Section 2 describes a scenario for explaining the system usage and system architecture of the proposed VCSR system, and Section 3 describes the method and results of the experiment. Discussion of the results is presented in Section 4, and the conclusion is drawn in Section 5.

VCSR System

In this section, a scenario is given to explain how the proposed VCSR system recommends new videos to users and what kind of role the system plays. Next, the system architecture is described in detail using three modules. Finally, the inherent differences among various video recommendation systems are compared in order to identify the functionality of the VCSR system.

Scenario: An Intermediate Role in a Digital Library

Usually, there are many video resources in a library. Whenever a new video arrives, the librarian needs to make a video introduction. One way that he or she can do this is to refer to the simple description on the video cover. The only other way to extract the necessary information from the video would be to watch it. It is important to note that information taken from a video cover may not be detailed enough to make an appropriate recommendation based on a learner’s interests.

![Figure 1. VCSR system workflow using a digital library as an example](image)
Using the proposed VCSR system, all of the above problems can be turned into an automated process. Figure 1 describes the automated workflow, which is mediated by the VCSR system. First, the library receives the ordered video resource regularly. (a) When a new video is incoming, the system will receive it as an input. (b) The system will then automatically produce a video summary. (c) The summary is then compared to profile information registered by users, and (d) users receive a multimedia-based recommendation email, composed of the video cover, a video description, online video clips, hyperlinks to the top 5 ranked video summaries, and the video summaries themselves, consisting of text and key frames that correspond to the text (see Figure 2). In the recommendation email, users can click the image frames and watch the video clips online.

Figure 2. An illustration of the recommendation email content

System Architecture

The system architecture of the VCSR system is illustrated in Figure 3. Once a new video is incoming, the Video OCR Module recognizes captions as video caption documents. These documents are then passed to the Summarization Module and the summary documents for the video are generated by extracting the key passages. Finally, the video recommendation emails are generated by the Recommendation Module, which estimates the relevance for each learner according to their profiles. By combining these three modules, the system can automatically generate and send video recommendation emails when there is a new incoming video. In other words, the entire process in the VCSR system is automated without any human intervention. Each of the three modules is described below.

Video OCR Module

The Video OCR Module processes the input video frame sequence and recognizes all the captions. Video images can provide rich visual information to people, but video speech content plays a more significant role in understanding video content. In many educational films, such as those produced by the Discovery Channel and National Geographic, rich caption information is an excellent way to describe the video content. Therefore, video captions are extracted as video speech content for further processing.
VCD and DVD are the most popular video formats in the market. Since captions are embedded in these kinds of video formats, it is easy to extract them from video resources. By employing video Optical Character Recognition (video OCR) techniques, these captions can be automatically extracted and identified. Using video OCR techniques in the VCD format is a more difficult task than in DVD format due to the fact that captions and image frames are not independent. Because most of the video resources used in this study are VCD films, we have employed OCR systems (Lee, Wu & Chang, 2005; Wu, Lee, Yang & Yen, 2006; Wu, Yang & Lee, 2008) for caption content extraction. As reported in a previous study (Wu et al., 2006), the performance of such video OCR systems was about 70%-80%. The Video OCR module first decomposes the input video into a list of frames, and then performs the following steps: (a) filtering to remove the noisy blocks; (b) representing the character vector; and (c) character recognizing, in order to find the most similar characters. Finally, the recognized words form the video caption document.

**Summarization Module**

The Summarization Module processes the video caption document from the Video OCR Module. Over the past few years, text summarization has been thoroughly studied by researchers and is well-developed. However, most of the traditional summary generation methods aim to extract a set of key sentences from documents, but this is not the case for video summarization. This is because the key sentences in the video are less meaningful in an educational context and usually not what users are interested in. To overcome these problems, the Q/A-based (Question & Answering) approach (Wu et al., 2006; Wu et al., 2008) is adopted, which is more likely to route the information that users want to know, thereby enhancing the motivation to learn.

In this module, the video caption document is initially divided into five segments based on the time sequence analysis. Passage-level summarization is more suitable than sentence-level summarization with regard to video summary generation because video contains multiple sub-topics (Lee et al., 2005). For each segment, therefore, the video Q/A-based approach is adopted to extract passage-level answers and form the video summary document. Here, it is assumed that using each segment’s answers as the video summary could be more complete and comprehensive in relation to understanding video content. In order to enhance the readability, the top 5 ranked video summaries are emailed to users. Finally, these summaries are stored in the video summary database for future browsing.

**Recommendation Module**

The Recommendation Module processes the video summary generated by the previous two modules. It then compares the video summary with the user profiles, which record each user’s personal information (name, email
address, interests, and major subject). An XML-based format is used to store each user’s data. Figure 4 illustrates a fragment of user profiles as an example. Users can edit their profile at anytime through an interface.

```xml
<profiles>
  <user id="VCSR_USER_0001">
    <firstname>王</firstname>
    <lastname>聰</lastname>
    <email>coral@cl.nctu.edu.tw</email>
    <interests>
      <item>歷史</item>
      <item>軍事</item>
      <item>地理</item>
    </interests>
    <majorSubject>運動學問科技</majorSubject>
    <favoriteMovie>Independence Day</favoriteMovie>
  </user>
  <user id="VCSR_USER_0002">
    <email>john@sample.com</email>
    <interests>
      <item>音樂</item>
      <item>電影</item>
      <item>美食</item>
    </interests>
    <majorSubject>藝術科學</majorSubject>
    <favoriteMovie>Forrest Gump</favoriteMovie>
  </user>
  <user id="VCSR_USER_0003">
    <email>sara@sample.com</email>
    <interests>
      <item>旅行</item>
      <item>攝影</item>
      <item>美術</item>
    </interests>
    <favoriteMovie>La La Land</favoriteMovie>
  </user>
  ...
</profiles>
```

*Figure 4. XML-based description of user profiles*

For the recommendation, the focus is on calculating the relevance of the video summary for users. In order to match users and information more effectively, the video title, video description, and video summary are integrated as the sources for matching profiles with relevant video information. If the video content and the user profile match, the recommendation module will extract the email address in the profile and send a recommendation email automatically. In order to achieve adaptive personalization, the system does not comprehensively notify users of all videos. Instead, the related videos are presented to users, preventing the system emails from becoming spam. When a new video is incoming, the recommendation module compares the content of this video with all users. By means of comparing the user profiles to the generated video summaries, only a portion of the users will receive the recommendation email. The key is the similarity measurement between the two information sources. The similarity measurement is estimated with a cosine value (Baeza-Yates & Ribeiro-Neto, 1999) and described as follows:

\[
Sim(UP_i, SUM_j) = \frac{\text{dot}(UP_i, SUM_j)}{\|UP_i\| \|SUM_j\|}
\]

Where \(\text{dot}(X, Y) = (X_1Y_1 + X_2Y_2 + ... + X_nY_n)\), \(\|X\|\) is the one-norm of the vector \(X\) and can be estimated as follows:

\[
\|X\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}
\]

Each \(x_i\) is a word in either the summary or the profile. \(Sim(X, Y)\) computes the similarity (i.e., cosine value) between the two variables \(X\) and \(Y\). The terms \(UP_i\) and \(SUM_j\) denote the user profile for person \(i\), and summary \(j\) for the \(j\)-th new videos. For each user, the following function is used to determine a relevant score with regard to a specific summary \(j\), \(SUM_j\):

\[
\text{RevScore}(i, SUM_j) = Sim(UP_i, SUM_j) - \theta
\]

where \(\theta\) is a pre-defined threshold. If \(\text{RevScore} > 0\), then the system considers the video as positive and relevant to the user, and it sends the recommendation email. Otherwise, the system skips the user. In other words, scores of a higher relevancy are more likely to be what the user is interested in. If the likelihood score exceeds the threshold, it will send the auto-generated video recommendation emails to the users.
The Differences between Video Recommendation Systems

A comparison of the differences between the proposed VCSR system and other video recommendation systems is shown in Table 1. The VCSR system can receive raw video data and extract captions with the video OCR module. However, other systems only take artificially structured (e.g., title, genre, and user ratings) or semi-structured (e.g., description, summary, and review) data as input. The dataset is then compared with user profiles to select the suitable video. The VCSR system can automatically create video summaries from captions as a compared dataset with the summarization module. In other words, speech content is taken into account in the VCSR system. By contrast, other systems lack speech content summaries and ignore the importance of speech content when choosing videos. Moreover, other systems do not have an automatic summarization function. The manual summaries of other systems are usually subjective and influenced by advertisements. With the VCSR system, automatic summaries describe the video content more objectively. Finally, the VCSR system can play video clips online, which is another advantage over other systems.

<table>
<thead>
<tr>
<th>System</th>
<th>Input data type</th>
<th>Compared dataset</th>
<th>Content summary</th>
<th>Video play function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MovieLens (<a href="http://movielens.umn.edu">http://movielens.umn.edu</a>)</td>
<td>Structured</td>
<td>User ratings</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Alspector et al., 1998; Cotter et al., 2000</td>
<td>Structured</td>
<td>Content feature</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Basu et al., 1998</td>
<td>Structured</td>
<td>User ratings &amp; Content feature</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Melville et al., 2002</td>
<td>Structured &amp; Semi-structured</td>
<td>User ratings &amp; Content feature</td>
<td>Yes (manual summary)</td>
<td>No</td>
</tr>
<tr>
<td>VCSR system</td>
<td>Structured &amp; Unstructured (raw video)</td>
<td>Content feature (speech content especially)</td>
<td>Yes (automatic summary)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1. Comparison of video recommendation systems

Experiment

Method

The purpose of the experiment is to examine whether or not the VCSR system can motivate learners to watch videos. The experiment was conducted to support the research hypothesis that the VCSR system can recommend adaptive videos, and that video summaries can increase the learner’s motivation to watch videos. In this experiment, thirty subjects (nineteen male and eleven female) were employed, all of whom were either undergraduate students or graduate students between the ages of 18 and 25 years old. They came from different colleges of the National Central University in Taiwan, but most of the subjects were students of the College of Electrical Engineering and Computer Science. One hundred and eighty-one Discovery films were selected as the video sources because Discovery video data is one of the most popular learning materials. Two email types are used to examine the hypothesis: text-based emails and multimedia-based email. The text-based email contains a video cover image and short descriptions (derived from the video cover), as shown in Figure 5a. The multimedia-based email contains a video cover image, short descriptions, the extracted summaries (roughly 25 sentences and the corresponding image frames), and video clips (see Figure 5b). The former is mainly used as an adjunct to the typical text-based recommendation, while the latter will help us examine the impact and the effectiveness the VCSR system in comparison with the text-based recommendation.

The instruments used in this experiment are questionnaires that were designed according to the strategies of the ARCS model. The features of the ARCS model are identified as: attention (A), which refers to the extent to which the learners’ attention is aroused; relevance (R), which refers to the learners’ perception about whether the content of the recommendation email is related to personal needs or past experience; confidence (C), which refers to the learners’ perceived likelihood of achieving their expected goal after using the system; and satisfaction (S), which refers to the system preference based on learners’ user experience. Thus, this study examines the effectiveness of the proposed system by using the ARCS model. The research questions are, “Would the system improve learners’ motivation?” and “Which function of the system improves motivation?”
The experimental procedure is illustrated in Figure 6. A two-phase study was designed in order to explore learning motivation with regard to summarization and recommendation. In order to understand the subjects’ video-watching habits and their personal information, they were first asked to fill out a habit questionnaire and a user profile which included their name, email address, and interests. Since the study has a within-subjects design, all subjects participated in both phases. Subjects could autonomously decide the experiment time by themselves, and the experiment was recorded by a camera for further analysis. In the first phase, subjects got text-based recommendation emails from the system. They were then asked to fill out the user experience questionnaire. Next, subjects took a rest before the beginning of the second phase. In the second phase, subjects got multimedia-based recommendation emails from the system.
emails from the system. Similarly, they were then asked to fill out the second questionnaire after using the system. Finally, subjects were interviewed in an effort to understand their preferences in relation to the two types of emails.

Results

The results of the questionnaires are presented in Table 2. The categories of the questionnaires are divided into four factors relating to motivation (attention, relevance, confidence, and satisfaction) and two factors relating to system functions (summarization and recommendation).

### Table 2. The results of the questionnaires.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Phase 1 Mean</th>
<th>SD</th>
<th>Phase 2 Mean</th>
<th>SD</th>
<th>Variation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>3.54</td>
<td>.92</td>
<td>4.17</td>
<td>.75</td>
<td>.62</td>
<td>.000***</td>
</tr>
<tr>
<td>Relevance</td>
<td>4.03</td>
<td>.56</td>
<td>4.48</td>
<td>.64</td>
<td>.45</td>
<td>.029*</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.03</td>
<td>.76</td>
<td>4.30</td>
<td>.65</td>
<td>.27</td>
<td>.013*</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>3.71</td>
<td>.71</td>
<td>4.12</td>
<td>.68</td>
<td>.41</td>
<td>.000***</td>
</tr>
<tr>
<td>Summarization</td>
<td>3.37</td>
<td>.96</td>
<td>4.10</td>
<td>.66</td>
<td>.73</td>
<td>.000***</td>
</tr>
<tr>
<td>Recommendation</td>
<td>3.87</td>
<td>.90</td>
<td>3.93</td>
<td>.87</td>
<td>.07</td>
<td>.313</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; ***p < .001.

Attention: This factor focused on whether the recommendation emails aroused learners’ attention. The results showed that the Phase 2 emails, on average (M = 4.17, SD = .75), received higher attention than the Phase 1 emails (M = 3.54, SD = .92). There is a significant difference between Phase 1 and Phase 2 with regard to the attention factor (p < .001).

Relevance: Relevance refers to whether or not the recommendation emails connected the learners’ prior experiences. The results showed that Phase 2 had a higher mean (M = 4.48, SD = .64) than Phase 1 (M = 4.03, SD = .56). There is a significant difference between the two phases with regard to the relevance factor (p < .05).

Confidence: Confidence means that the content provided by the recommendation emails can enhance learners’ confidence in their ability to develop their interest. Results in Phase 1 (M = 4.03, SD = .76) and Phase 2 (M = 4.30, SD = .65) were both positive, yet there is a significant difference between Phase 1 and Phase 2 (p < .05).

Satisfaction: Satisfaction refers to the user experience. The results showed that the difference between Phase 1 and Phase 2 is very significant (p < .001). 82% of subjects in Phase 2 strongly agreed that the system was easy to use (M = 4.12, SD = .68).

Summarization: Summarization refers to the system’s ability to provide sufficient information for users to understand the main idea of the recommended video. The findings indicate that the summarization in Phase 2 (M = 4.10, SD = .66) was significantly superior to that of Phase 1 (M = 3.37, SD = .96). Again, there is a significant difference between the two phases (p < .001).

Recommendation: This factor is concerned with the accuracy of the recommendation. Because the same recommendation mechanism is employed in both phases of the experiment, it is not a surprise that there was no significant difference in the results.

Discussion

Multimedia content can support learning by helping learners to identify emergent goals within a context. This system aims to motivate learners and provide them with a useful tool for knowledge assimilation. As the results show, learners received the information they sought and assimilated knowledge unknowingly through watching the videos. Emails from Phase 2 proved to be more useful in helping learners assimilate knowledge than the emails from Phase
1, owing to the versatility of the VCSR system. The obvious differences between text-based email and multimedia-based email are the summary and the interactive video play. According to the analysis of the recorded data from the camera, subjects spent about one minute and thirty seconds reading the text-based emails in Phase 1, and they spent more than five minutes in Phase 2. The multimedia-based emails, therefore, motivated subjects more successfully. A large number of subjects reported favorably on the effectiveness of the summaries and video clips. The main reason they wanted to watch the video was because of the detailed summary and video clips. They could grasp the meaning of video through the abstract summary instead of watching the whole video, and they could then determine exactly which video to watch.

According to the results of the questionnaires, the system promotes more motivation with regard to the attention factor because the system actively sends content-rich emails to learners. When learners checked their mailboxes, they saw the video content of the recommendation email. If the video content was free, learners would read the summary and then watch the video content. On the other hand, if the video content was not free, learners could read the summary to get main idea of the video content. Furthermore, it was found that the relevance and confidence factors are related to both the accuracy of the recommendation and the abundance of summarization. Because subjects felt that the recommendation email was related to their past experience and the summary offered rich information, they had confidence in developing their interests. Subjects gave positive feedback in the two phases in relation to the recommendation mechanism, which results in a significant difference between the two phases with regard to summarization. Subjects expressed that the system increased video-watching frequency and extended their knowledge. Furthermore, they hoped to own the system in the future, which accounts for the large difference between phases in relation to the satisfaction factor.

In addition, results also showed that subjects with different backgrounds had different perceptions. Some subjects answered that they had not been interested in watching videos in the past, but after operating the system, they expressed a keen interest in watching videos. They were excited at the prospect of easily receiving video information and exploring a previously unknown area of knowledge. Some subjects highly praised the recommendation function and video clips of the system, saying that they really needed a tool to help them find new multimedia materials and confirm their video choices. In the past, they had spent a lot of time searching for multimedia data, only knowing the content of a video after watching it. By using the system, they spent more time watching relevant videos than searching and watching irrelevant media.

The original purpose of summarization is to design an objective and faithful summary that is not influenced by advertisements. The results of the experiment show that the system successfully achieved this goal. Subjects fully and quickly understood the content of the recommended videos. One subject, however, expressed that he would like to get some subjective comments about the video from others. He suggested that the system set up a forum to give feedback and form discussion groups with other system users. Although the system did not provide this function, it is important to note that this subject shifted from a passive role to an active one, wanting to participate in discussions about the videos, rather than simply using the system and accepting or rejecting recommendations. In addition, a few subjects mentioned that there were some irrelevant words in the summaries. This phenomenon results from OCR errors. Furthermore, some subjects said that they would prefer to get more information in the summary if the video contains rich educational content or knowledge they are interested in. However, if the genre of a video is commercial entertainment, the subjects expressed a lack of interest in receiving summary information, saying that more summary information would reduce their imagination as it relates to the story. In this experiment, of course, all of the materials are educational in nature, and the summaries give learners a clearer understanding of the content of the videos. Learners are also motivated to watch the entire video in an effort to gain more knowledge with respect to the topic of the summary.

Generally, the more information learners get, the more motivation is generated. If learners receive more detailed information, there is a better chance that their interest will be piqued. However, the fact that more recommendation information might inundate learners with redundant information is certainly a concern. Some subjects might treat the recommendation email as junk, thereby lowering their interest in receiving future recommendation emails from the system. On the contrary, most subjects expressed that they had dreaded searching videos in the past, and that the system considerably reduced the effort required to find the desired media. One subject expected to have a function that allowed him to save and manage recommendation emails. This indicates that the subject was pleased with the system. He wanted to keep the information and have his own accessible store of summaries and recommendations. In addition, some subjects suggested that the recommendation categories need to be diversified. For example, if
someone likes animals, but especially insects, and the system recommends a video about dolphins, he will surely reject it. Moreover, subjects suggested that the system needs to analyze the learners’ personality. If someone is curious about a certain topic, then the system should give him a specific part of the suitable video. It is important to note that the employed video data is limited and may not cover all subjects. Therefore, there is little likelihood that the system did not find a suitable video to match subjects’ tastes in this experiment. This might be why the recommendation scores in both phases are so close: the system reinforced the subjects’ knowledge, but it could not change an uninteresting item into an interesting one. Some subjects also pointed out that the recommendation accuracy of the system is very important. If most of the content in the recommendation email refers to learners’ specific interests, learners will be pleasantly surprised and have higher a confidence level. On the contrary, if the system gives an inaccurate recommendation, learners would slowly lose faith in the accuracy of the system.

The VCSR system can extract a summary from numerous unstructured data (raw video), in turn constructing a recommendation. Meanwhile, the summary content was compared to learners’ profiles in order to generate recommendation emails. One subject expressed that he preferred multimedia emails because the detailed summaries. He said he would feel no preference for either of the two email types if the multimedia-based email did not include a summary. Hence, the recommendation email is not fascinating for the user if it lacks a summary. Another subject said that he was not interested in Discovery films, so he did not pay particular attention to the summary. The recommendation mechanism must necessarily conform to learners’ interests. If the recommendation mechanism is not robust, learners are less likely to read the summary. The above observations indicate that summarization and recommendation are intimately linked in the sense that the summary is important to the perceived validity of the recommendation.

**Conclusion**

In this paper, we presented an automatic multimedia content summarization and adaptable recommendation system, called VCSR, which is able to auto-recommend suitable video content for learners. The system can automatically create a video summary from raw video data and send it to learners based on their user profiles. The results of the experiment demonstrated that the system positively promoted learners’ motivation in comparison with text-based recommendation method. Furthermore, the experiment suggests that there could be a relationship between summarization and recommendation in the sense that if there is no summary in the recommendation email, the recommendation email is likely not appealing enough to motivate learners. On the other hand, learners do not like to read the summary if the recommendation mechanism is not robust enough to fit their needs.

In the future, we plan to collect the personal information of more learners. The more the system understands learners’ personalities, the better the recommendations it can create. Based on each learner profile, we also plan to customize the video summary for each learner. This implies that different learners can obtain different summaries even though the video source is the same. In addition, we are going to create a forum where learners can give their feedback and discuss issues relating to the VCSR system. Learners will play an active role rather than a passive role in the system. Moreover, it is necessary to improve the accuracy of the OCR module and the quality of the summaries by employing a more effective algorithm. Finally, we expect that the VCSR system could be applied to a real educational setting, such as a digital library, in the near future.

**Acknowledgements**

The authors would like to thank Mr. Chia-Hsing Shen for assisting in the experiment. The authors would also like to thank all the subjects who participated and cooperated in the experiment. This study was partially supported by a grant (NSC 97-2628-S-008-001-MY3) from the National Science Council of Taiwan.

**References**


A Study on the Learning Efficiency of Multimedia-Presented, Computer-Based Science Information

Ying-Hua Guan
Department of Applied Chinese Languages and Literature, National Taiwan Normal University, Taiwan // yhguan@ntnu.edu.tw

ABSTRACT
This study investigated the effects of multimedia presentations on the efficiency of learning scientific information (i.e. information on basic anatomy of human brains and their functions, the definition of cognitive psychology, and the structure of human memory). Experiment 1 investigated whether the modality effect could be observed when the learning material contained auditory information and visuals altered in complexity, and whether the redundancy effect is caused by redundant information or by interference in information processing. In Experiment 2, verbal-only information was used to examine whether subjects could perform better with auditory rather than with on-screen textual information, and whether the length of the verbal information would exert an effect on learning. The results of Experiment 1 contradicted the prediction of the modality effect in that subjects learned no better or even worse with the audio-visual format of learning material than did subjects with the visual-only one. Besides, redundant information per se did not impair learning, which suggested that the redundancy effect could be rather caused by the interference in information processing. The results of Experiment 2 indicated a negative effect of auditory information on learning regardless of the length of the verbal information. No evidence supported the superiority of auditory instructional mode over the visual one.

Keywords
Multimedia presentations, Learning efficiency, Modality effect, Working memory, Cognitive load theory

Introduction
Multimedia presentations are widely used in e-learning environments. Learning material is often presented with text, audio, video, and static pictures, whereby same information is sometimes repeatedly presented by different media. According to the cognitive load theory, when the learning material has high intrinsic load or high element interactivity and when they are presented as text with any kind of visuals, it is better to present texts aurally because the efficiency of information processing in working memory will be enhanced if information is presented in different modalities (i.e. the modality effect). When visuals are combined with text presented visually, learning is impaired because learners have to split their attention between text and visuals, in order to integrate both sources of information (i.e. the split-attention effect). The modality effect and the split-attention effect seem to be repeatedly confirmed in a series of studies conducted by Sweller, Mayer and their co-workers (Mousavi, Low and Sweller, 1995; Sweller, van Merriënboer, and Paas, 1998; Kalyuga, Chandler, and Sweller, 1999; Mayer and Moreno, 1998, 2002; Moreno and Mayer, 1999; Mayer, Moreno, Boire, and Vagge, 1999).

Advocates of cognitive load theory often refer to the studies reviewed by Penney (1989) and the working memory model proposed by Baddeley (Baddeley and Hitch, 1974). However, when examining what exactly those reviewed studies and Baddeley’s model indicated, it can be found that the theories claiming the positive effect of dual mode presentation on learning are highly dubious. First, the studies reviewed by Penney actually investigated the nature of short-term retention of verbal stimuli. Typical tasks for the subjects of those studies consisted of the recall of a series of digits or single words. It was found that the auditory stimuli almost always resulted in higher recall than did the visual ones, which was termed the modality effect. Some researchers plead for the superiority of auditory information over the visual one based on their belief in the modality effect. The problem is that remembering digits or a word list greatly differs from normal learning situations in which no such short-term memory but rather the comprehension of the learning material is required. There is actually no valid theoretical basis to apply the observed phenomena in those studies to multimedia learning theory. As Penney already put it in her article “In spite of the large and robust effects of presentation modality found in short-term memory tasks, there was no evidence of any permanent effects on learning, and modality effects in long-term memory tasks were conspicuously absent.” (Penney, 1989: 398).
Second, when pictorial material is involved, some theorists referred to Baddeley's working memory model and Paivio's dual coding theory (Paivio, 1986, 1991) to back up their theories. Again, both theories suggested that working memory contains modality- or code-specific processors, but the theories did not indicate that working memory always has more capacity to process information presented in different modalities than if it were presented in the same modality. According to Baddeley's working memory model (Baddeley and Hitch, 1974; Baddeley 2000), working memory consists of a central executive and at least three subsidiary systems — the phonological loop, the visuo-spatial sketchpad, and the episodic buffer (see Figure 1). The central executive is an attentional controller that supervises and coordinates the subsidiary systems in working memory. It was characterized as “a limited capacity pool of general processing resources”. The phonological loop has the function of processing speech or printed text, whereas the visuo-spatial sketchpad is responsible for setting up and manipulating mental images. The episodic buffer is assumed to be the place where information from the subsystems of working memory and that from long term memory is integrated. According to Baddeley, “It is assumed to be episodic in the sense that it holds integrated episodes or scenes and to be a buffer in providing a limited capacity interface between systems using different codes... This allows multiple sources of information to be considered simultaneously, creating a model of the environment that may be manipulated to solve problems and plan future behavior” (Baddeley, 2001: 858).

![Figure 1. The working memory model by Baddeley (2000: 421)](image)

According to the cognitive theory of multimedia learning (Moreno and Mayer, 2000), when learners simultaneously process pictorial and verbal information, they need to integrate both types of information in the working memory in order to form a mental model based on their understanding of the learning material (see Figure 2). In Baddeley’s model, it is the central executive or the episodic buffer that is responsible for integrating information from different subsystems. However, the capacity of the central executive or the episodic buffer is very limited and will not be substantially increased even if the phonological loop and the visuo-spatial sketchpad are simultaneously used to process information.

![Figure 2. Cognitive theory of multimedia learning (Moreno & Mayer, 2000).](image)
Furthermore, it is not plausible not to associate a split-attention effect with dual mode presentation because the general processing resources of the central executive or episodic buffer need to be shared or divided for processing two types of information at the same time, which certainly leads to a split of attention.

Recently, more and more studies provided results imposing limitations on the modality effect. Tindall-Ford, Chandler, and Sweller (1997) found that the modality effect was only obtained for learning materials with high element interactivity but not for those with low element interactivity. Jeung, Chandler, and Sweller (1997) found that using auditory texts did not improve learning when they were presented together with visually high-demanding diagrams. Tabbers, Martens, and van Merriënboer (2004) even found a reversed modality effect when students were asked to study multimedia-presented learning material on a computer for an hour and were tested subsequently. They concluded that the reversed effect was probably attributed to the relatively long learning time and the self-controlled learning speed. Guan (2003, 2006) found negative effects of dual mode presentations on learning efficiency in a series of multimedia learning scenarios in which subjects were allowed to learn the instructions at their own paces without restricted learning time. Moreover, Ginns (2005) conducted a meta-analysis of the modality effect based on research review and found that the pacing of presentation determines the occurrence of the modality effect. Under system-paced conditions in which the learning time and the playing of audio instructional material were controlled by the experimenters, the modality effect occurred, whereas the opposite effect occurred under self-paced conditions in which learners could control their own learning speed and were not limited in terms of learning time and the frequency of playing of audio instructional material. All the studies mentioned above suggest the need of re-examining the validity of the modality effect proposed by the cognitive load theory and the cognitive theory of multimedia learning.

In terms of the redundancy effect, Kalyuga (Kalyuga, Chandler, & Sweller, 1999, 2000; Kalyuga 2000) found that presenting texts aurally was not beneficial for learning when the same texts were already presented visually, or when the information given in the auditory texts was not useful to the more experienced learners. This phenomenon refers to the redundancy effect. Since redundant information can often be found in multimedia-based learning material, it is necessary to investigate whether learning is impaired simply because some of the information was repeated in different formats or because the presentation mode actually caused interference in information processing. It is assumed that simultaneously presenting the same text visually and aurally can easily cause interference between reading and listening to the text because the speed of reading is usually faster than that of listening. Hence, it is intriguing to examine whether the interference would be eliminated when the speed of presenting visual and auditory text is synchronized. If learning is not impaired in this case, it follows that it is not the redundancy of information that impairs learning but the interference elicited by the simultaneous processing visual and auditory text.

In line with the issues discussed above, this study aimed to investigate whether using dual mode presentation (i.e. audio-visual presentation) could really facilitate the learning of scientific information, whether presenting the same information in two different modalities really impaired learning, whether subjects could learn better with auditory text than with on-screen text, and whether the length of text played a role in this regard. In this study, two experiments were carried out to investigate these issues. Experiment 1 examined the validity of the modality effect and the redundancy effect in the context of learning science information. Experiment 2 investigated whether the modality effect existed for verbal-only information in different lengths. Self-paced learning was adopted for both experiments because it resembles the real e-learning situations more closely.

**Experiment 1**

The goal of this experiment was to investigate whether using dual-mode presentation really facilitates the learning of scientific information, especially whether the modality effect could always be observed when the learning material contained auditory information with visuals of different complexities, and finally, whether the redundancy effect is simply caused by the presence of redundant information or rather by the interference during information processing.

**Materials**

Nine different multimedia presentations were used to present the instructional material with the same content about the basic anatomy of human brain and its functions. The nine conditions are as follows: Condition 1 [on-screen
text/simple diagram], Condition 2 [on-screen text/medium diagram], Condition 3 [on-screen text/complex diagram],
Condition 4 [auditory text/simple diagram], Condition 5 [auditory text/medium diagram], and Condition 6 [auditory
text/complex diagram]. In Condition 7 [auditory text/animation], animation was applied to complex diagram as a
visual aid in order to examine whether animation could help reduce the load of visual search. Condition 8 [running
on-screen text/animation] was adapted from Condition 7 by using system-paced running text instead of the auditory
one. Condition 9 [running on-screen text + auditory text/animation] employed system-paced running text and
auditory text together with animation. The content of the instructional material consisted of four sections, which was
displayed by seven Web pages: (1) the structure of the brain, (2) specific functions of cerebral cortices, (3) the
functions of the midbrain and the limbic system, and (4) the coordination and integration of central nervous system.
Examples of the instructional material are given in figures 3-6.

Figure 3. An example of the instructional material for
Condition 1

Figure 4. An example of the instructional material for
Condition 2

Figure 5. An example of the instructional material for
Condition 3

Figure 6. An example of the instructional material for
Condition 8

Subjects and Procedures

The subjects were 178 students at a university in Taiwan, whose majors were Applied Linguistics, Physics, Business
Administration, or Electrical Engineering. Firstly, they were asked to take a pretest which consisted of 12 multiple
choice questions concerning the basic anatomy of human brain and its functions. The test questions were constructed
in line with the three criteria of Bloom’s taxonomy: knowledge, comprehension, and evaluation. The knowledge
questions tested subjects’ memory of the facts that were explicitly stated in the learning material (e.g. What is the
function of primary sensory cortex?). The comprehension questions, on the other hand, tested whether subjects could
translate knowledge into new context and predict consequences (e.g. If a patient with brain lesion can speak fluently,
but his speech is meaningless, and he cannot understand other people’s speech, where could the lesion site locate in
his brain?). The evaluation questions tested whether subjects could compare and discriminate between ideas (e.g.
Which of the following is not true in terms of the functions of basal ganglia and limbic system?).

Subsequently, the subjects were randomly assigned to the experimental conditions and were asked to learn the
computer-based instructional material at their own pace. While they were learning, their learning time, viewing
sequences as well as viewing frequencies were registered by a computer. Although the learning time was not limited,
subjects were told that they should concentrate on learning and their learning time would be taken into account for
the evaluation of their learning efficiency. A subject could finish learning if he or she had the confidence to take the
posttest. To achieve the consistency and reliability of pretest and posttest, the same test was used for both cases, though the subjects were not aware of this before they took the posttest.

The measurement of learning efficiency

If subjects’ learning times were not taken into account, the scores of posttest would be the direct measurement of their learning outcomes. However, based on the belief that efficient multimedia presentations should enable learners to process information easily and quickly, I suggest that the learning time should be considered as well. Parallel to my previous studies (Guan, 2003), learning efficiency in this study was calculated as follows:

\[ \text{Learning efficiency} = \frac{\text{score of posttest}}{\text{learning time (in seconds)}} \times 100 \]

To avoid comparing learning efficiency in fractions, it is multiplied by 100. According to this formula, it followed that learning efficiency was high when the score of posttest was high while the learning time was short.

Hypotheses

If the modality effect occurs, subjects who received auditory information with simple or medium complexity of diagrams should outperform their counterparts, but not if the diagrams are too complex. Using animation as guidance for visual search should reduce the load on visual memory and thus, restore the modality effect (cf. the study by Jeung et al., 1997). Finally, if the redundancy effect exists, presenting the same text both visually and aurally along with animation should greatly impair learning.

Results

The data of some subjects were excluded because they either failed to finish viewing all instructional material or to complete the posttest. The mean pretest scores, mean posttest-scores, the standard deviations, and the number of subjects of Conditions 1 to 6 are given in Table 1. A Levene’s test for homogeneity of variances was computed for the pretest scores to examine whether the variances of subjects’ performance on pretest did not significantly differ from each other across conditions. The result showed that the difference of variances was not significant (F(8, 136)=1.453, p=0.18). For conditions 1 to 6, an ANCOVA was computed for two main factors text mode (visual vs. auditory) and picture complexity (simple, medium, complex) with pretest score as the covariate. The result showed that the effect of text mode on the posttest scores was significant (F(1, 102)=4.434, p<0.05, partial $\eta^2 =0.042$), whereas the effects of picture complexity and the interaction between the two main factors were not significant. The subjects receiving learning material with on-screen text (mean=6.492, SE=0.216) performed significantly better than their counterparts who received learning material with auditory text (mean=5.822, SE=0.234). Moreover, posthoc tests (Bonferroni) showed no significant difference in posttest scores between the six experimental conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean pretest score</th>
<th>Standard deviation</th>
<th>Mean posttest score</th>
<th>Standard deviation</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: on-screen text/simple static</td>
<td>3.29</td>
<td>1.68</td>
<td>6.19</td>
<td>1.29</td>
<td>21</td>
</tr>
<tr>
<td>2: on-screen text/medium diagram</td>
<td>3.74</td>
<td>1.76</td>
<td>6.79</td>
<td>1.93</td>
<td>19</td>
</tr>
<tr>
<td>3: on-screen text/complex diagram</td>
<td>3.89</td>
<td>1.76</td>
<td>6.53</td>
<td>1.47</td>
<td>19</td>
</tr>
<tr>
<td>4: auditory text/simple diagram</td>
<td>3.81</td>
<td>1.52</td>
<td>5.63</td>
<td>2.31</td>
<td>16</td>
</tr>
<tr>
<td>5: auditory text/medium diagram</td>
<td>3.11</td>
<td>1.84</td>
<td>5.72</td>
<td>1.60</td>
<td>18</td>
</tr>
<tr>
<td>6: auditory text/complex diagram</td>
<td>3.94</td>
<td>2.11</td>
<td>6.12</td>
<td>2.19</td>
<td>16</td>
</tr>
</tbody>
</table>

Nonetheless, when comparing the learning efficiency, significant differences were found between the experimental conditions. Table 2 shows the mean learning efficiency, the standard errors, and the number of subjects of each experimental condition.
Table 2: The mean learning efficiency of Conditions 1 to 6

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean learning efficiency</th>
<th>Standard Error</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: on-screen text/simple static</td>
<td>9.403</td>
<td>0.593</td>
<td>21</td>
</tr>
<tr>
<td>2: on-screen text/medium diagram</td>
<td>9.632</td>
<td>0.621</td>
<td>19</td>
</tr>
<tr>
<td>3: on-screen text/complex diagram</td>
<td>8.310</td>
<td>0.622</td>
<td>19</td>
</tr>
<tr>
<td>4: auditory text/simple diagram</td>
<td>6.524</td>
<td>0.677</td>
<td>16</td>
</tr>
<tr>
<td>5: auditory text/medium diagram</td>
<td>6.833</td>
<td>0.642</td>
<td>18</td>
</tr>
<tr>
<td>6: auditory text/complex diagram</td>
<td>6.196</td>
<td>0.678</td>
<td>16</td>
</tr>
</tbody>
</table>

The results of ANCOVA showed that the effect of text mode on learning efficiency was significant ($F(1, 102) = 24.868, p < 0.001$, partial $\eta^2 = 0.196$). The subjects receiving on-screen text (mean=9.125, SE=0.351) learned much more efficiently than their counterparts who received auditory text (mean=6.53, SE=0.381). The effects of picture complexity and the interaction between the two main factors did not reach statistical significance. Posthoc tests (Bonferroni) indicated significant differences between Conditions 1 and 4, Conditions 1 and 6, Conditions 2 and 4, Conditions 2 and 5, Conditions 2 and 6.

In order to find out whether using animation could reduce the load of visual search and create a certain modality effect, the performance of subjects in Conditions 6 and 7 was compared. According to ANCOVA, the use of animation did not exert an effect on the learning efficiency of those subjects. The subjects receiving the complex diagrams (n=16, mean=6.429, SE=0.504) performed slightly better than those receiving animated complex diagrams (n=10, mean=5.445, SE=0.637), but the difference was not significant.

Furthermore, for examining whether the redundancy effect was simply caused by presenting redundant information, the performance of subjects in Conditions 8 and 9 was compared. The result of ANCOVA showed that the effect of using redundant information on learning efficiency was statistically not significant but reached a certain practical significance ($F(1, 23)=2.489, p=0.13$, partial $\eta^2 = 0.098$, effect size: medium. According to Cohen (1973), Thompson (2006) and Vacha-Haase & Thompson (2004), the effect size is medium when partial $\eta^2$ is larger than 0.06 and smaller than 0.14). The subjects in Condition 8 (n=14, mean=6.483, SE=0.548) tended to perform worse than the subjects in Condition 9 (n=12, mean=7.793, SE=0.594), which contradicted the redundancy effect.

Finally, subjects’ performance on answering the three types of questions – knowledge, comprehension, and evaluation – was analyzed. A series of ANCOVA analyses with the main factor instruction mode and pretest scores as the covariate were computed. The results indicated that the effect of instruction mode on the scores of knowledge was significant ($F(8, 133)=2.936, p< 0.01$, partial $\eta^2 = 0.15$). Posthoc test (LSD) showed that the subjects of Condition 4 achieved much lower scores on knowledge questions than did the subjects of other conditions. (mean=2.129, SE=0.23, mean accuracy rate: 42.58%). The mean accuracy rate was calculated by dividing the mean score on knowledge by the total score on knowledge. Subsequent mean accuracy rates followed the same pattern. No significant difference was found among subjects of other conditions and the overall mean score of knowledge was 3.05 with the mean accuracy rate of 61%. Moreover, the effect of instruction mode on the scores of comprehension was significant ($F(8, 133)=2.63, p<0.01$, partial $\eta^2 =0.139$), whereas that on the scores of evaluation was not. The overall mean score of comprehension was 2.39 with the mean accuracy rate of 59.64%, and that of evaluation was 0.92 with the mean accuracy rate of 30.7%. Posthoc test (LSD) showed that the subjects of Condition 9 achieved significantly higher scores on comprehension questions (mean=3.325, SE=0.277, mean accuracy rate: 83.13%) than did the subjects of other conditions.

Discussion

In this experiment, subjects were allowed to learn the information without a time limit. Since the learning time was not equal across the subjects, I suggest that it is more meaningful to compare the learning efficiency rather than comparing the posttest scores. For one, the effects of different instruction modes on learning could not be revealed because subjects could spend longer time to compensate for the weakness of certain instruction modes. For another, a truly beneficial instruction mode should enable learners to learn efficiently, i.e. achieve the learning goal with as little learning time as possible. As the results demonstrated, when the learning time was not considered, the posttest scores alone were not able to distinguish subjects’ learning performance across the conditions. It is therefore
legitimate and more meaningful to compare the learning efficiency instead of only comparing the posttest scores. In the following, I shall focus on discussing the results in regard to learning efficiency.

The results of Experiment 1 contradicted the modality effect in that the subjects who received auditory instructions combined with simple or medium complexity of diagrams performed significantly worse than the subjects who received the same diagrams combined with on-screen text instructions. When the diagrams were complex, no significant difference was found in the performance between subjects receiving on-screen or auditory text. With regard to the effect of animation, the result indicated that the use of flash movie did not successfully assist visual search and did not reduce the load of visual memory because the subjects of Condition 7 performed no better than those of Condition 6.

Furthermore, it is interesting that the subjects of Condition 9 tended to achieve higher learning efficiency than did the subjects of Condition 8. This result indicated that the redundancy effect was not simply caused by presenting redundant information to the learners. It is assumed that the speed of running on-screen text was synchronized with that of the auditory text, and as a consequence, it was not difficult to follow the instructional information. By contrast, if the on-screen text was not running, or if the speed of both (visual and auditory) texts was not synchronized, strong interference would impair the processing of information. Nevertheless, how subjects actually processed the information of animation, running on-screen text and auditory text simultaneously, and how they benefited from these presentations, remain unclear. The possibility that the synchronized presentation of running on-screen text and the auditory text might enhance memory cannot be ruled out. Further investigations are certainly required to clarify this issue.

Finally, subjects’ performance on the three types of questions indicated that subjects’ performance on knowledge questions was the best, which was only slightly better than that on comprehension questions but substantially better than that on evaluation questions. It should be noted that subjects’ performance on comprehension items was very close to that on knowledge, which clearly demonstrated that the subjects did not just memorize the content of the learning material but also achieved deeper understanding of the content, i.e. a higher level of learning. The most interesting result was that the subjects of Condition 9 outperformed all other subjects in comprehension questions. It seems that using system-paced running on-screen text with synchronized auditory text and animation helped subjects achieve deeper understanding of the learning material. The subjects of Condition 4 did not perform well on knowledge questions. This could be due to the fact that information provided in simple diagrams was either sparse or more scattered. Thus, it might be difficult for the subjects to integrate the pictorial information with the auditory information during learning.

**Experiment 2**

The objective of Experiment 2 was to investigate whether subjects could perform better with auditory rather than with visual information when verbal-only material was provided, and whether the length of verbal information would exert an effect on learning.

**Materials**

The learning material contained two topics: one was about the definition of cognitive psychology, which consisted of 182 Chinese characters, whereas the other one was concerned with the structure of human memory, which comprised 607 Chinese characters. Each topic was presented by a single web page. Two presentation modes were employed: auditory versus visual. Examples of the long and the short instructional material presented by on-screen text are shown in Figures 7 and 8.

**Subjects and procedures**

The subjects were 165 students who also participated in Experiment 1. They were randomly assigned to one of the experimental conditions. Due to the amount of headphones available in the computer lab, more subjects were recruited for the on-screen condition (n=93) than for the auditory one (n=72). The subjects were asked to learn the
information of both topics that were presented by one of the presentation modes. They were allowed to learn at their own pace with the learning time registered by computer. They took pretests and posttests for assessing their levels of knowledge before and after learning the information. For the condition with longer information both pre- and posttest consisted of eight multiple choice questions, whereas for the condition with shorter information both tests contained three multiple choice questions. The questions of those tests were based on two criteria of Bloom’s taxonomy: knowledge and evaluation. Examples of knowledge questions for long and short instructional condition are as follows: “What is the primary form of encoding in short-term memory?”, and “What is cognition?”, whereas the examples of evaluation questions for each condition are: “Which of the following has nothing to do with amnesia?”, and “Which of the following is not studied in the field of cognitive psychology?”. 

![Figure 7](image1.png) Figure 7. The instructional material on the definition of cognitive psychology.

![Figure 8](image2.png) Figure 8. The instructional material on the structure of human memory.

**Results**

For the long instructional conditions, the data of some subjects were excluded because the subjects failed to finish their posttest. Table 3 showed the mean pretest and posttest scores, the standard deviations, and the number of subjects of on-screen text and auditory text conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean pretest score</th>
<th>Standard deviation</th>
<th>Mean posttest score</th>
<th>Standard deviation</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>on-screen text</td>
<td>4.45</td>
<td>1.37</td>
<td>6.65</td>
<td>1.10</td>
<td>92</td>
</tr>
<tr>
<td>auditory text</td>
<td>4.20</td>
<td>1.45</td>
<td>6.02</td>
<td>1.21</td>
<td>65</td>
</tr>
</tbody>
</table>

To make sure that the unequal number of subjects did not bias the results, a Levene’s test for homogeneity of variances was computed for the pretest scores. The result indicated that the variances of subjects’ performance on pretest of both conditions were roughly the same (F(1, 155)=0.107; p=0.744). An ANCOVA was then computed for the main factor presentation mode (on-screen vs. auditory) with the pretest score as the covariate. The result showed that the effect of presentation mode on the posttest scores was significant (F(1, 154)=10.435, p<0.005, partial $\eta^2$ =0.063). The subjects who learned with on-screen text outperformed their counterparts.

Moreover, when comparing subjects’ learning efficiency between the two conditions, the result indicated that the effect of presentation mode on the learning efficiency was significant (F(1, 154) = 47.80, p < 0.001, partial $\eta^2$ =0.237). The subjects who learned with on-screen text performed significantly better than the subjects who learned with auditory text. The mean learning efficiency, standard error and number of subjects of the two conditions are given in Table 4.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean learning efficiency</th>
<th>Standard Error</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>on-screen text</td>
<td>5.97</td>
<td>0.266</td>
<td>92</td>
</tr>
<tr>
<td>auditory text</td>
<td>3.10</td>
<td>0.317</td>
<td>65</td>
</tr>
</tbody>
</table>
The analysis of subjects’ performance on the two types of questions showed that the effect of presentation mode on knowledge was significant (F(1, 153)=15.477, p<0.001, partial $\eta^2 =0.092$) while that on evaluation was not. The subjects learned with visual text (mean=5.311, SE=0.092, mean accuracy rate: 88.52%) outperformed their counterparts (mean=4.749, SE=0.109, mean accuracy rate: 79.15%) in terms of answering knowledge questions.

For the short instructional conditions, the data of 10 subjects of on-screen text condition were further excluded because they also failed to finish the posttest. Table 5 shows the mean pretest score, the mean posttest score, the standard deviations, and the number of subjects of the on-screen and auditory text conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean pretest score</th>
<th>Standard deviation</th>
<th>Mean posttest score</th>
<th>Standard deviation</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>on-screen text</td>
<td>1.18</td>
<td>0.61</td>
<td>1.76</td>
<td>0.56</td>
<td>82</td>
</tr>
<tr>
<td>auditory text</td>
<td>1.05</td>
<td>0.65</td>
<td>1.86</td>
<td>0.46</td>
<td>65</td>
</tr>
</tbody>
</table>

According to the Levene’s test for homogeneity of variances, the variance of subjects’ performance on pretest of both conditions did not differ significantly (F(1, 145)=0.67; p=0.414). An ANCOVA was computed for the main factor presentation mode (on-screen vs. auditory) with the pretest score as the covariate. The result showed that the effect of presentation mode on the posttest scores was not significant. However, when comparing the learning efficiency, the result of ANCOVA indicated that the effect of presentation mode on the learning efficiency was significant (F(1, 144) =20.85 , p <0.001, partial $\eta^2 =0.126$). The subjects who learned with on-screen text significantly outperformed the subjects who learned with auditory text. Table 6 shows the mean learning efficiency, standard errors and number of subjects of each condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean learning efficiency</th>
<th>Standard Error</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>on-screen text</td>
<td>4.22</td>
<td>0.283</td>
<td>82</td>
</tr>
<tr>
<td>auditory text</td>
<td>2.27</td>
<td>0.318</td>
<td>65</td>
</tr>
</tbody>
</table>

Subjects’ performance on answering the two types of questions indicated that presentation mode did not have any effect on their performance on knowledge or evaluation. The overall mean score of knowledge was 0.77 with a mean accuracy rate of 38.25%, and the mean score of evaluation was 1 with a mean accuracy rate of 50%.

Discussion

The results of Experiment 2 clearly showed that the auditory presentation mode of text did not produce any superiority in learning over the on-screen presentation mode, regardless of the length of text. Instead, negative effects were constantly found with the auditory text. In line with these results, the alleged superiority of auditory text was not confirmed in this study.

Subjects’ performance on the two question-types indicated that for the long instruction conditions, subjects learning with visual text could more correctly recall the facts about the learning material in comparison to the subjects learning with auditory text. This is clear evidence against the modality effect proposed by the cognitive load theory. For the short instructional conditions, however, no significant difference was found between subjects’ performance on answering the two types of questions, which indicated that even short auditory text could not exert any positive effect on recalling or judging the facts given in the text.

General Discussion and Conclusion

In this study, the modality effect and the redundancy effect were examined. The results of both experiments showed no evidence supporting the benefit of using dual mode presentation for instructional material. Instead, counter-evidence was repeatedly found when comparing subjects’ posttest scores and learning efficiency: dual mode
presentation yielded either negative or no effect on subjects’ posttest scores and learning efficiency. These findings seem to be consistent with those of the studies allowing self-paced learning with unlimited learning time. As Tabbers et al. (2004) found a reversed modality effect in their study, they concluded that the reversed effect was probably due to the relatively long learning time and the self-controlled learning speed. In line with this, the results of the current study along with those of my previous research and the studies reviewed earlier indicate that the modality effect is probably only observable under highly controlled experimental conditions in which learners are asked to learn a short piece of information with a limited fixed learning time. When the learning material is longer, and when students can learn the instructional material at their own pace, a negative effect can readily be found with dual mode presentation. Another important finding of this study was that the modality effect observed in the short-term verbal memory studies cannot be legitimately applied to multimedia-based learning scenarios in which subjects are required to comprehend and learn meaningful subject matter.

In terms of the redundancy effect, it was found that the presence of redundant information alone did not impair learning. Since the speed of self-running on-screen text was synchronized with that of the auditory text, it seems that the processing of visual and auditory text did not interfere with each other but was harmonized or even reinforced. Further studies need to be done to examine more exactly how learners process redundant and non-redundant information at the same time. The overall results of this study suggest that educational practitioners need to think seriously about the effects of multimedia presentations on learning efficiency. Especially, the use of dual-mode presentation as an instructional format should be considered cautiously, or it may only impair learning.

While the present study reveals interesting results, some limitations may exist in this study. For one, the subjects were all Taiwanese first-year university students with different majors that are not related to neuroscience, cognitive science or cognitive psychology. Whether the same result patterns could be replicated with students whose majors are related to the subject matter to be learned is unclear. For another, all subjects in this study speak Mandarin, a logographic language, which is different from alphabetic languages. It is important for future studies to further investigate whether the interaction between the simultaneous processing of auditory and visual verbal information would be different depending on the nature of the language.

Acknowledgement

Funding of this research is supported by National Science Council (Grant number NSC94-2511-S-033-003), Taiwan.

References


Designing a Syntax-Based Retrieval System for Supporting Language Learning

Nai-Lung Tsao¹, Chin-Hwa Kuo², David Wible¹ and Tsung-Fu Hung²

¹Graduate Institute of Learning and Instruction, National Central University, Taiwan, No.300, Jung-da Rd. JhongLi, Taiwan // Tel: +886-3-4227151-33851 // Fax: 886-3-4273371
²Department of Computer Science and Information Engineering, Tamkang University, Taiwan, No.151 Ying-chuan Road, Tamsui, Taiwan // Tel : +886-2-26215656-3304
beaktsao@gmail.com // chkuo@mail.tku.edu.tw // wible45@yahoo.com // kidd@mail.iwillnow.org

ABSTRACT

In this paper, we propose a syntax-based text retrieval system for on-line language learning and use a fast regular expression search engine as its main component. Regular expression searches provide more scalable querying and search results than keyword-based searches. However, without a well-designed index scheme, the execution time of regular expression search would be unacceptable to users. Our methods are based on Cho and Rajagopalan (2002) and we introduce some modifications, such as a presuf index constructing algorithm and a method for deciding minimum filter factor, to meet the requirements of our syntax-based text retrieval system. The experiment results show the index space size is small and the performance of syntax-based sample queries show significant improvements over benchmark results. A user-friendly query generator is designed to support users who have no background knowledge of regular expressions.

Keywords

Regular expression, Syntax-based retrieval, Indexing techniques, Query processing, Language learning

Introduction

The need for a well-design search engine is dramatically increasing because of the growing amount of data in real world, such as that on web pages, in standard text corpora and in movie databases. In addition to effective searching from massive amounts of data, recent search engines feature a flexible query language providing a wider variety of targeted search items. Compared to traditional text retrieval systems using keywords as basic query symbols, a system which provides regular expression as one of query languages seems more suited to meeting this requirement.

The main purpose of the syntax-based text retrieval system is to support grammatical querying of tagged corpora for language learners and teachers. While syntax-based queries contribute improvements to general-purpose queries of massive amounts of data, the power of regular expressions provide even further advantages when the users are language learners or teachers and their purpose is finding examples of specific types of language message. Consider the following example.

Example 1: In ESL teaching, it is important to teach learners the particular forms the certain verbs require of their complements, for example, keep requires a following verb in its –ing form (keep trying but not keep to try). If we want to find the form in any text corpora, traditional text retrieval systems have no good solutions but might scan all documents which include the keyword keep in the collection. Variant searching results are found, such as keep in touch and noun phrases followed by keep. That is because of the approaches to index construction, usually using inverted indices as index structure, in traditional text retrieval systems are extracting useful keywords first and using these keywords as indices. A regular expression can easily present such a pattern. The regex (regular expression) below presents one possible way to describe this pattern.

```
keep\s+\w+ing
```

Even though regular expressions provide more flexible querying, they still create a serious problem in terms of search response time. For example, a text collection with 1 million documents and 1,000 words average length would take an unacceptable response time, say, a couple hours, by match the above sample query pattern to strings of text. Without further processing, the only way to find the pattern is scanning each document one by one in the text collection.
There have been several proposals made to solve this search time problem. Most of them use k-gram index construction and build efficient index structures for quick searching of index terms. The index terms extracted from a text collection would be every sequence of characters of length k in each data unit. Once the index terms are extracted, the systems can construct suffix trees by using the technique described by Baeza-Yates and Gonnet (1996) or inverted indices in (Baeza-Yates and Navarro, 2004) (the most commonly used as index structure for k-gram indexing) to identify the data unit positions of each index term. By using the index, the search engine can only scan the documents or any data units which contain the specific targeted string in the regular expression query. For example, if a system performs k-gram index extraction from k=3-10, then the query in Example 1 can be matched just in data unit positions of index terms “keep” and “ing” rather than in the whole text collection. This simple idea can substantially reduce the search time. Every system would decide different ranges of k for specific purposes or considerations, such as limited secondary memory size. In this paper, the main approaches to building regex search engine are based on (Cho and Rajagopalan, 2002). The approaches use minimum index storage space and provide short enough search response time for most regex queries.

Example 1 can be implemented in any regex enabled search engine. What if, however, we want to find the syntax without any specific string, such as ing? This would be valuable for learners who do not yet know what form requirements a particular verb places on its complement and want to discover this through examples. Example 2 shows another syntax pattern commonly targeted in ESL teaching.

Example 2: Collocation is a persistent area of difficulty for ESL learners. For example, what verbs can be used before the noun problem? If the corpus is part-of-speech tagged, we can use the following regex to search.

\[<w \text{POSagrams}="V\W+">\w+</w> (<w[^>]+></w>){0,5} <w[^>]+>problem</w>\]

where POSagrams="V\W+" indicates words tagged to verb.

Of course the text collection needs further processing to make the query work and we will discuss this below. However, we can see after adding some word class information to each word, such as part-of speech, the power of regex query can be greatly expanded by indicating specific strings of words or word classes or some combination of the two.

There are some different aspects of designing syntax-based text retrieval system from regex search engine. Cho and Rajagopalan (2002) provide a hardware I/O consideration for deciding the filtering factor, say, 0.1, while they assume the speed of sequential I/O access is ten times faster than random access. In this paper, we propose a different way to decide this factor concerning syntax-based text retrieval. The details of the approach are described in “Syntax-based text retrieval system” section.

The rest of the paper is organized as follows: in Section “Regular expression search engine”, the main approaches provided by Cho and Rajagopalan (2002) are described. The design details of syntax-based retrieval system are proposed in following section, including a user-friendly query generator. The conclusion and future works are presented in the last section.

Regular expression search engine

Figure 1 shows the main components of a typical regular expression search engine. The parts most different from traditional keyword-based search engines are the index construction and query processing components. The following subsections focus on these two components, which are proposed by Cho and Rajagopalan (2002).

Index construction

This subsection deals with the index construction algorithm and structure. First we introduce some notations and definitions. A k-gram is a string \( x = x_1x_2\cdots x_k \), where \( x_i : 1 \leq i \leq k \) is a character. A data unit means the unit in which the raw data is partitioned, such as web page, a paragraph or a sentence in documents. Let \( M(x) \) denote the
number of data units which contain x and then the filter factor is denoted by $\text{ff}(x) = 1 - M(x)/N$, where N means the total number of data units. Now we can set the minimum filter factor $\text{minff}$ and only keep index terms with filter factors greater than $\text{minff}$. This would make the number of index terms smaller and more useful. For example, $\text{minff} = 0.95$ means the system only keeps the index term which can filter 95% of data units. This filter scheme can make the index terms more discriminative, which is important for a retrieval system. We call these index terms useful indices.

![Diagram of a typical regular expression search engine](image)

*Figure 1*: The illustrative figure of a typical regular expression search engine

Even if the system only maintains useful indices, the number of indices is still large. Every string expanded from a useful index will be useful, too. For example, if the index NBA is useful with in the text “How to buy NBA tickets”, then “y NBA” and “NBA t” are useful and it seems not necessary. Therefore, the system only maintains a presuf (prefix and suffix) free index set. A presuf free set means there is no x in the index set is a prefix or suffix of any other index $x'$. For example, {ab, ac, abc} and {bc, ac, abc} are not presuf free because ab is a prefix of abc in first set and bc is a suffix of abc in second set.

After determining index terms, we can construct the index for a regex search engine. We use inverted indices as our index storage structure, which is easily accessed by RDBMS. The algorithm is shown as follow:

```
Input : text collection
Output : index

[1] k=1, Useless={.;} // . is a zero-length string
[2] while (Useless is not empty)
[3]     k-grams := all k-grams in text collection whose (k-1)-prefix∈ Useless or (k-1)-suffix∈ Useless
[4]     Useless := {}    
[5]     For each x in k-grams
[6]         If ff(x) ≥ minff Then
[7]             insert(x,index)  //the gram is useful
[8]         Else
[9]             Useless := Useless ∪ {x}
[10]     k := k+1
```
In (Cho and Rajagopalan, 2002), the prefix-free indices are built first and then the suffix-free indices are considered. We reduce the number of passes needed for index construction from two passes to one pass and get speed improvement. The different part of our steps from the original one in (Cho and Rajagopalan, 2002) is we do prefix checking and suffix checking in the same pass. We need to scan entire data once for extracting all k-grams. After the index is built, it still needs one more entire data scanning for identifying index term positions. Therefore, the whole index construction needs only two entire data scanings without any extra memory space. However, if the memory is large enough, we can extract k-grams and get the positions in the same data scanning pass.

**Query processing**

As mentioned before, the k-grams index is used to reduce the number of data units to be matched by regex. For this reason, the query processor has to determine which index term to look up. The following shows the algorithm for determining candidate data units for regex matching.

**Input**: a regular expression query, \( r \)

**Output**: candidate data units

1. **Rewrite regular expression** \( r \) so that it only uses $, (, ), | and \&, where $ means entire data units and \& means AND operator. The details are as follow:
   1. Based on the order in Table 1, perform the action of target symbol in \( r \).
   2. If there is no operator in any two operands, insert \&
2. **Transform the infix expression to postfix expression**
3. **Get k-gram which has the smallest position set contained in operand string.**
4. **Perform the set operation among the position sets extracted in Step 3 and then get the final candidate data units for regex string matching.**

In Step 1, the query processor refers to Table 1 to eliminate useless symbols for identifying index terms and replace the useless symbols with operators or specific operands. For example, because any string between [ and ] presents only one character and almost every unigrams are not discriminative, the query processor will eliminate every [] and the string between [ and ], and use a & mark to replace AND operator. For example, after being processed by Step 1, the regex in Example 2 will become <w & POS=& >& <w &> (<w &> & <w &>) <w & >problem< & w>.

In Step 2, the query processor transforms the infix expressions to postfix expressions thereby eliminating the need for operator and bracketing priorities. For example, A&(B|C) becomes ABC|&

Because not every k-gram is in index, the string operands in final postfix expression would not be available in index. Therefore, a strategy has to be provided to identify which index term to look up. We use a simple strategy: get k-gram which has the smallest position set. For example, an operand is Compaq and only two indexed k-grams contained in Compaq are comp and paq. The number of position set of comp is larger than the number of position set of paq. On this account, the query processor picks the position set of paq for generating candidate data units for regex string matching. This method is Step 3 in the algorithm.

**Table 1**: Regular expression symbols and handled actions

<table>
<thead>
<tr>
<th>Description</th>
<th>Action</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>\   Escape character((\d) is any numeric characters )</td>
<td>Delete this symbol and the next character</td>
<td>1</td>
</tr>
<tr>
<td>[]  Any character in the bracket([a</td>
<td>b</td>
<td>c]=a</td>
</tr>
<tr>
<td>+   One or more repetition of previous character</td>
<td>Delete this symbol</td>
<td>3</td>
</tr>
<tr>
<td>‧   Any character(a.d match asd)</td>
<td>Delete this symbol</td>
<td>4</td>
</tr>
<tr>
<td>*   Zero or more repetition of previous character</td>
<td>Delete this symbol and replace with $</td>
<td>5</td>
</tr>
<tr>
<td>^   Any character except the next character</td>
<td>Delete this symbol and next token if next token is not blank</td>
<td>6</td>
</tr>
</tbody>
</table>
Another thing that should be noticed is what happens if an operand contains no index terms in the index. Because of the useful index design, we cannot treat this situation as empty candidate set. Unfortunately, this operand must be treated as if all data units might contain this string. Based on this problem mentioned above, another problem raised here is after the position set operation, the candidate data units might be the whole data units. Two kinds of strategies might be used: (1) recommend users to change the query and (2) only list first n results for users. We will adopt the second method in our system.

Syntax-based text retrieval system

As far as we know, there have been few researches about syntax-based search. Corley et al. (2001) proposed a syntactic structure search engine. Their system selects sentences by some grammar rules. This is achieved by a fast chart parser. Compared to their approach, our system needs no grammar rules or parsers, but linear syntax knowledge. Our query language is simple and intuitive.

In this section we will discuss the design of syntax-based text retrieval systems. The index construction and query processing are just like the descriptions in previous section. We will focus on corpus design and evaluating sample queries.

Tagged corpus

As mentioned above, this syntax-based retrieval system is designed for language learning. We want to provide a system which can search grammatical information in a target text collection. Therefore the raw text has to be preprocessed for advanced searching. According to the grammar tutoring samples from high school textbooks, we decide to add part-of-speech information to the raw text. We train a Markov Model-based POS tagger (2000) and use British National Corpus (BNC, http://www.natcorp.ox.ac.uk/) as our training data. The internal evaluation shows this tagger has 93% precision including identifying unknown words.

The information of the tagged corpus also contains lemma of each word. The lemmatizing operation should be performed after POS tagging because some words have different lemma with different POS. The following shows an example after POS tagging and XML structure formatting.

```xml
<s>
  <w Lemma="do" POS="VDD"> Did </w>
  <w Lemma="you" POS="PNP">you</w>
  <w Lemma="know" POS="VVI">know</w>
  <w Lemma="?" POS="PUN">?</w>
</s>
```

We use XML to wrap the tagged text for three reasons. First, XML is suitable for representing grammatical information. Basically, XML and English can all be treated as a tree by visualization (Allen, 1995). Although our corpus only has POS information now, we think other grammatical information, such as phrase structure and relative clause identification, can easily be represented by XML. Second, many XML parsers have been produced. Using these XML parsers can easily extract information encoded by XML node inside and outside. Lastly, XML has been standard for data exchange. This will prepare our data to run on other applications some day.

Filter factor

We have discussed how Cho and Rajagopalan(2002) decide their minimum filter factor. This idea is not suitable for our system. Since our target users are English teachers and learners, they have more need of English vocabulary searching capacity but not special formats, such as URL or telephone numbers. Under this assumption, the index terms in our system should be token-oriented. Further, because length of most of English words is more than 3 and those words of length 1 or 2, such as I, is, and to, lack discriminative value for retrieval, we set our k from 3 to 10 (From the result of extracting experimental corpus, we found out there are about 3500 occurrences for 11-grams,
which we think it’s not worth to save the index.). As for the length 1 and 2 words, the system builds the isolated keyword index for them. While meeting length 1 and 2 words in query processing, the process will search this isolated index. Furthermore, because we want all words of length 3 kept in the index set (to make the searching results nonempty), the index construction starts on k=4. We also think minff should be increased as k is increased because it’s obvious that when the index length grows up, it can filter out more data. Therefore we change the function to \( \text{minff}(k) = 1 - \alpha \frac{k_0}{k-1} \times \max M(x)/N \) where \( k_0 \) means the smallest number in k, in our case, \( k_0 = 4 \). \( \alpha \) is a real number smaller than 1. \( \alpha \) can be \( 1/52 \) if no matter which letter of the English alphabet is added in front of or in back of the string with length k-1, each of them has the same \( M(x) \).

**Evaluation**

We collect some sample syntax-based regex query results to evaluate the system performance. There are numerous queries which are interesting for high school teachers or relevant to textbook content. The following shows ten query samples.

1. **Adj+respect**  
   
   \(<[^>]*SOS="\{0,1\}"[^>]*/[^<]>s*[^>]*/[^<]="N[S\{0,2\}[^<]">respect</w>"

2. **Verb+pain**  
   
   \(<[^>]*SOS="\{0,2\}"[^>]*/[^<]>s*([^>]*)[^<]>s*([^>]*)[^<]>s*([^>]*)[^<]">pain</w>"

3. **break +PREP**  
   
   \(<[^>]+break\w+</w>\s*[^>]*/[^<]="PR[S\{0,1\}[^<]">to</w>"

4. **of + WH-Word**  
   
   \(<[^>|>]+of</w>\s*[^>]*/[^<]="(PNQ|DTQ)"[^<]>[^<]+</w>"

5. **approach + to + Ving**  
   
   \(<[^>]+approach\w<[^<]/>s*[^>]*/[^<]="\w+ing</w>"

6. **on the other hand**  
   
   \(<[^>]+on</w>\s*[^>]*/[^<]="\w+ing</w>"

7. **no matter**  
   
   \(<[^>]+no</w>\s*[^>]*/[^<]="\w+ing</w>"

8. **as+Adj+as**  
   
   \(<[^>]+as</w>\s*[^>]*/[^<]="AJ[S\{0,1\}[^<]">as</w>"

9. **worth+Ving**  
   
   \(<[^>]+worth</w>\s*[^>]*/[^<]="VVG[^<]+</w>"

10. **to+Ving**  
    
    \(<[^>]+to</w>\s*[^>]*/[^<]="VVG[^<]+</w>"

The testing database is a subcorpus from BNC. We collect about 1,750,000 sentences, about 22,000,000 words, from BNC. All sentences, or say, data units, are POS tagged, lemmatized and XML formatted. The index construction approaches we use differ in some respects from (Cho and Rajagopalan, 2002). We do not extract k-grams from a sentence, but from keywords because we think the phrase searching can also be achieved by this method and for the above reason, this method can reduce the index storage. Table 2 shows the index sizes and posting sizes of complete index construction and presuf free index construction.
Table 3 shows the number of each candidate and final result number of each query and Figure 2 shows the execution time of the ten query samples by entire data search and our regex search engine. Most of the queries show a remarkable improvement by using our regex search engine. Those which show no improvement are due to the regex strings containing no useful index term. These unsuccessful queries contain no substring of keywords in sentences but POS tags strings. Even if the system treats this POS tags as index terms, these index terms almost all refer to entire data units or half of them. We treat this problem as one of our future works.

<table>
<thead>
<tr>
<th>Table 2: Sizes of index terms and postings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Index Terms</strong></td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Numbers of Postings</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3: Numbers of candidate set and final results. Total number: 1,777,516</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Query</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

**Figure 2**: Execution time of each query

**Query Generator**

Although the syntax-based retrieval system yields impressive performance gains, it is not easy to use for our target users, English teachers and learners. For this reason, we also have designed a query generator for those users who have no knowledge of regular expressions. The query generator is shown in Figure 3. Since our query language is linear, users construct a query element by element. Here an element means one grammatical unit, such as a word, a lemma, or a word class.
As shown in Figure 3, the user interface consists of three parts, namely, (1) keyword, (2) POS, and (3) anyword to construct the desired syntax form. The resulting format is shown in query input box (see legend 4 in Figure 3.) Note that the corresponding regex form can appear as users press the “Add this” button. Also note that the operation sequence of the above parts depends on the desired syntax format. In other words, it does not necessarily follow the order shown in Figure 3.

We make use of the following examples to illustrate the above description.

**Example 3**, single word: To retrieve sentences contain the keyword wonderful, user simply type in wonderful in the keyword part, and press add button. Then the query input is appeared in query input box.

**Example 4**, syntax format: To retrieve the syntax format:

“verb keep followed by -ing form verb”

First we type in keep as first query term (legend 1 in Figure 3). Since the -ing word may be allocated at several words right from the word keep, we can insert words for this elasticity of demand (legend 1 in Figure 3). In this case, we add 0 to 5 “Anyword” patterns, which means –ing word is far away from keep at most 5 words. The last query term, of course, is the part-of-speech “-ing form verb” (legend 2 in Figure 3). The corresponding syntax query is generated (keep + <anyword>(0,5) + Ving) as legend 4 in Figure 3.

**Example 5**, syntax format: To retrieve the syntax format:

“adjective before the word rain”

First, we choose the adj term from POS term. Then, add any other words (e.g., from zero to three words). Finally, we type in the word rain in the keyword term. The corresponding syntax query is generated (adj + <anyword>(0,3) + rain).

![Figure 3: The designed User Interface of Query Generator](image-url)
Conclusion

We have described how to design a syntax-based text retrieval system for language learning. The system introduces a regular expression search engine to reduce response time. Grammatical annotation of the target texts helps the system to provide syntax-based searches to locate tokens of specific types of target language use, thus providing valuable supports for language teaching and learning. However, there is still no solution for making class-based queries faster because they have no index available to reduce the search time. Enrichment of the corpus information, for example with semantic tagging, is also one important work for future research.

Acknowledgment

The work described in this paper was partially supported by the grants from the National Science Council, Taiwan (Project Nos. NSC 95-2520-S-032-004-MY3 and NSC 96-2524-S-008-003-). 

References


A Knowledge-based Approach to Retrieving Teaching Materials for Context-aware Learning

Wen-Chung Shih
Department of Information Science and Applications and Department of Computer Science and Information Engineering, Asia University, Taichung 41354, Taiwan // wjshih@asia.edu.tw

Shian-Shyong Tseng*
Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan // sstseng@cs.nctu.edu.tw // Department of Information Science and Applications, Asia University, Taichung 41354, Taiwan // sstseng@asia.edu.tw

*Corresponding author

ABSTRACT
With the rapid development of wireless communication and sensor technologies, ubiquitous learning has become a promising solution to educational problems. In context-aware ubiquitous learning environments, it is required that learning content is retrieved according to environmental contexts, such as learners’ location. Also, a learning content retrieval scheme should be able to work with various instructional strategies for different learning activities. To solve the context-aware learning content retrieval problem, we propose a strategy-driven approach to derive content retrieval strategies from instructional strategies. Moreover, we construct a knowledge-based system to expand query keywords based on the derived strategies and then select relevant keywords according to geographical distance between entities of concept and learners. This system is composed of four components: knowledge transformation, query expansion, content retrieval and user interface. Besides, ontology construction algorithms, designed for teachers to easily build up ontology from course outlines, are applied to generate the rules of query expansion and the taxonomic index of learning object repository. The experimental results indicate that the proposed approach can increase the learning performance of students.

Keywords
Ubiquitous learning, Context-aware learning, SCORM, Information retrieval, Ontology, Knowledge-base

Introduction
With the rapid development of wireless communication and sensor technologies, ubiquitous learning (u-learning) has become a promising solution to educational problems which can sense the situation of learners and provide adaptive supports to students (Chen et al., 2007; Hwang et al., 2008; Kuo et al., 2007; Si et al., 2006; Yang, 2006). Context-awareness is the major characteristic of u-learning, and the situation or environment of a learner situated can be sensed. There are twofold advantages of context-aware learning. One can alleviate environmental limitations, and the other can utilize available resources to facilitate learning.

There are several types of applications for context-aware u-learning. A typical scenario is “learning with on-line guidance” as presented in Hwang’s research (2006), which focused on the “identification of plants” unit of the Nature Science course in an elementary school. The context is in campus, and the human-system interaction is described as follows:

- System: Can you identify the plant in front of you?
- Student: Yes.
- System: What is the name of this plant?
- Student: Ring-cupped oak.
- System: Do you see any insect on it?
- Student: Yes.
- System: Can you identify this insect?
- Student: No.
- …
The assumption is that the system is aware of the location of the student and her/his nearby plants by sensor technologies and built-in campus maps.

Learning activities in ubiquitous environments are directed by instructional strategies which are general approaches instead of specific methods. As shown in Figure 1, instructional activities are generated according to instructional strategies originated from pedagogic theories. Designers of learning activities should utilize the advantages of u-learning environments to realize pedagogic goals.

![Layered relation of instructional activities, strategies and theories](image)

The context-aware learning content retrieval problem is motivated by the following assumptions:

- Students’ learning performance can be improved by providing right content at right time and right place.
- During ubiquitous learning, students’ queries are usually related to knowledge of their nearby objects.

Retrieval of learning content, hereafter named Content Retrieval (CR), is an important activity in u-learning, especially for on-line data searching and cooperative problem solving. Furthermore, both teachers and students need to retrieve learning content for teaching and learning respectively. However, conventional keyword-based content retrieval schemes do not take context information into consideration, and therefore they cannot fulfill the basic requirements of u-learning to provide users with adaptive results. To support context-aware learning, learning contents need to be provided according to learners’ contexts. For example, when a student can not identify an insect in the u-learning course, s/he can access a learning object repository for more information by submitting a query. Thus we can imagine that queries are most likely ambiguous and need refinement. If context information can be applied to refine the original query, it will be easier for learners to retrieve relevant contents.

As shown in Figure 2, we classify the schemes of content retrieval into static and dynamic types according to the adaptability of the retrieved results. For static CR, the retrieved result only depends on the query regardless of users and contexts. Nevertheless, dynamic CR can be further divided into personalized, context-aware, and other schemes according to the factors considered by the adaptive mechanisms of CR. The static CR is adapted to subjective factors of learners, such as user profile, preference, etc. In other words, the same query submitted by different persons could result in different results retrieved. On the other hand, context-aware CR is adapted to objective factors of learners, like time, place, device, activity, and peers etc. Hence, the same query run in different contexts could obtain distinct results.

Learning content retrieval is a universal requirement for many learning scenarios, such as Intelligent Tutoring Systems, and Zone of Proximal Development etc. However, each scenario has its own needs for content retrieval. In particular, an important characteristic of context-aware ubiquitous learning is to provide right contents to learners at right place and right time. In other words, it is required for a retrieval system to acquire the contents adapted to the learners’ contexts.
In this research, we investigate the context-aware learning content retrieval problem, which is to retrieve relevant learning contents from a repository according to the given query and context information, to improve the precision and recall of retrieval. However, there are some difficulties in the process of our research. First, context information needs to be taken into account for context-aware retrieval. Therefore, traditional information retrieval schemes have to be enhanced as context-aware. Second, needs for teaching materials are related to pedagogic factors, such as instructional strategies and goals. It is required to design a retrieval scheme which is flexible enough to adapt to various instructional strategies. Third, characteristics of standardized learning content must be considered to improve the accuracy of similarity comparison, such as metadata and structural information. Accordingly, the acquisition of context information requires extensive deployment of sensors. In this paper, we assume that the module of context acquisition is available and focus on the previous three issues.

To overcome the aforementioned difficulties, we propose a strategy-driven approach enabled by a knowledge-based query expansion method. First, we intend to expand the original query by acquired context information to retrieve content which is adapted to learners’ contextual environments. We adopt the technique of query expansion because most queries in web search are short, ambiguous, and refinement-needed. Second, we propose a knowledge-based approach to expand queries based on instructional strategies. According to our observation on ontology, such as Wordnet, basic strategies of query expansion include generalization, specialization, association, and their combination etc. For example, when the educator aims to encourage the learners to do high-level thinking, it may be appropriate to adopt an expansion technique of generalization which offers more general keywords for content retrieval. In this study, we assume that the content about entities near the learner is more relevant than that far from the learner. For example, when walking by a fern plant, we may want to find some content introducing the fern. Also, this work assumes the instructional strategy and the strategy mapping are defined by experts in advance. Otherwise, that will be our future work and we may focus on applying retrieval strategies to realize context-aware content retrieval.

Based on the aforementioned idea, we designed a system consisting of four components: knowledge transformation, query expansion, content retrieval and user interface. In the knowledge transformation component, algorithms of ontology building and rule generation are proposed, and therefore teachers can easily construct ontology from course outlines. The purposes of the ontology are to generate rules of query expansion and to construct taxonomic index of learning object repository. Besides, the remaining three components work as follows. In user interface, the user submits a query, and the context information is extracted by sensors. Next, in query expansion component, candidate keywords are inferred for query expansion, and keywords with entities closer to the learner are selected. Finally, in content retrieval component, results are retrieved according to the expanded query, and they are ranked by a similarity function considering characteristics of learning content. To speed up the searching process, we use a taxonomic index, which is generated by reorganizing the existing documents based on a bottom-up approach (Shih et al., 2008).

We think the proposed context-aware retrieval method can benefit the ubiquitous learning scenario by providing suitable content adapted to learners’ context and instructors’ strategies. Experiments have been conducted to show the evidence of this claim. First, an experiment involving 20 elementary school students is conducted to show the

---

**Figure 2. Classification of Content Retrieval**

- **Static Content Retrieval**
- **Dynamic (Adaptive) Content Retrieval**
- **Personalized Content Retrieval**
- **Context-aware Content Retrieval**
learning performance affected by the proposed retrieval method. Next, a survey involving 15 elementary school students is performed to understand their degree of satisfaction for this system. The results show that this system can streamline the retrieval process and facilitate the learning activity as well. In addition, the comments of teachers indicate that this system can effectively find suitable contents adapted to context and instructional strategies.

The contributions of this paper can be summarized as follows. First, we propose a strategy-driven approach to solve the context-aware learning content retrieval problem. This new approach integrates pedagogic requirements and technical solutions, and thus it can benefit both parties. Second, a knowledge-based system is designed to support query expansion, which can increase maintainability. Moreover, the flexibility of the knowledge-based approach facilitates future integration with educational strategies. In addition, the distance-based keyword selection can achieve context-awareness. Third, knowledge transformation algorithms are designed for automatic derivation of ontology and query expansion rules, and thus it can resolve the difficulties for teachers to manually construct ontology and code rules. Finally, we build up a prototype and obtain experimental results to show that this approach can increase accuracy, and it is helpful to context-aware learning.

**Preliminaries and Related Work**

In this section, we review background knowledge about u-learning, Sharable Content Object Reference Model (SCORM), context-aware information retrieval and query expansion. Moreover, related researches about this work are surveyed.

### Ubiquitous Learning

The rising of u-learning results from the convergence of e-learning and ubiquitous computing. However, this topic is too new to get a well accepted definition. Hwang et al. (2008) in their paper compared u-learning systems with m-learning systems and proposed twelve models for u-learning activities. Illustrative examples in that paper help readers understand what u-learning is like. In Si et al.’s paper (2006), existing u-learning applications were categorized as shown in Table 1, and a frame-based model was proposed to represent context-aware applications.

<table>
<thead>
<tr>
<th>Application Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location-aware learning guidance</td>
<td>Museum guide (Oppermann &amp; Specht, 1999)</td>
</tr>
<tr>
<td></td>
<td>Tour guide (Abowd et al., 1997)</td>
</tr>
<tr>
<td></td>
<td>Conference assistant (Dey &amp; Futakawa, 1999)</td>
</tr>
<tr>
<td>Correlation-aware collaborative learning</td>
<td>Japanese polite teaching (Yin et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Knowledge awareness map (El-Bishouty &amp; Ogata, 2006)</td>
</tr>
<tr>
<td></td>
<td>P2P content access and group discussion (Yang, 2006)</td>
</tr>
<tr>
<td>Task-aware supervised learning</td>
<td>Requirement satisfied learning (Cheng et al., 2006)</td>
</tr>
</tbody>
</table>

### Sharable Content Object Reference Model (SCORM)

To share and reuse teaching materials, several standards have been proposed recently. Among those standards, SCORM (http://www.adlnet.org/) is the most popular standard for learning contents sharing. It was proposed by the U.S. Department of Defense’s Advanced Distributed Learning (ADL) organization in 1997. This standard consists of several specifications developed by IEEE LTSC (Learning Technology Standards Committee, http://ltsc.ieee.org/wg12/), IMS (Instructional Management System, http://www imsproject.org/), and AICC (Aviation Industry CBT Committee, http://www.aicc.org/) etc. The SCORM specifications are a composite of several specifications developed by international standards organizations. In a nutshell, SCORM is a set of specifications for developing, packaging and delivering high-quality education, and for training materials whenever and wherever they are needed (Nitto et al., 2006). In SCORM, content packaging scheme is proposed to package the learning objects into standard teaching materials. The content packaging scheme defines a teaching materials package consisting of four components: 1) Metadata, which describes the characteristics or attributes of this learning content; 2) Organizations, which describe the structure of the teaching material; 3) Resources, which denote the
physical files linked by each learning object within the teaching material; and 4) the (Sub) Manifest, which describes this teaching material, consisting of itself and other teaching materials. SCORM Metadata refers to the IEEE’s Learning Object Metadata (LOM), and describes the attributes of teaching materials. IEEE LOM v1.0 includes nine categories: General, LifeCycle, Meta-Metadata, technical, educational, rights, relation, annotation, and classification.

**Information Retrieval**

Inverted file indexing has been widely used in information retrieval (Baeza-Yates & Ribeiro-Neto, 1999). An inverted file is used for indexing a document collection to expedite the searching process. The structure of an inverted file consists of two components: the vocabulary and the posting list. The vocabulary is composed of all distinct terms in the document collection. For each term, a list of all the documents containing this term is stored. The set of all these lists is called the posting list. However, the structure of a document is not counted in this model.

Storage requirements of inverted indices (Lee et al., 1996) have been evaluated based on B+-tree and posting list. Five strategies of the index term replication were discussed. This approach is extended to analyze the storage requirement of the proposed approach in this paper. In Cambazoglu & Aykanat’s paper (2006), 11 different implementations of ranking-based text retrieval systems using inverted indices were presented, and their time complexities were also investigated.

**Query Expansion**

A number of researches have focused on query expansion and tried to solve the ambiguity problem of short queries. Cui et al. (2003) divided a work on automatic query expansion into two classes: global analysis and local analysis. In Bhogal et al.’s paper (2007), query expansion approaches were reviewed and classified into three categories: relevance feedback, corpus dependent knowledge models and corpus independent knowledge models. Actions of query expansion can be based on various ideas. The point is how to determine the correlation between each pair of keywords. Relevance feedback uses top-ranked items in the returned set. Corpus independent models make decision by thesaurus. Corpus dependent models calculate co-occurrence in the corpus. For educational purposes, teacher-oriented schemes are needed. Our idea is based on geographical proximity. The recommended keyword is mainly related to the learner’s location instead of the original query.

There were few researches focused on query expansion for education. Kumela et al. (2004) used a query expansion technique implement advanced search features in support of Constructivist Learning. They claimed that learners need a mechanism to retrieve learning materials to support Constructivist Learning. Lee et al. (2008) presented an ontological approach to retrieve learning objects. They proposed an ontology-based query expansion algorithm for inferring user intention according to the original query. Experiments were conducted in terms of Precision, Recall and F-measure. These two researches indicate that applying query expansion to the search of educational resources is promising. However, they did not address the learning performance brought by the advanced search technology.

**Related Work**

SCORM-compliant teaching materials can be seen as structured documents. For fast retrieval of information from structured documents, Ko et al. (2002) proposed a new index structure integrating element-based and attribute-based structure information to represent a document and presented three retrieval methods: top-down, bottom-up and hybrid methods. Although this index structure takes information of elements and attributes into account, it is not suitable for management of a huge amount of documents due to its time and storage complexity. There have been numerous studies on Structured Document Retrieval (Trotman, 2004, 2005). Researches showed that structured searching can increase precision. Previous work mainly addresses XML and SGML documents. Besides, XML Query Languages, such as XIRQL, XQL, etc., were proposed. However, intra-document structural modeling is not suitable for SCORM-compliant documents.

For sharing and reusing teaching materials in different e-learning system, the Sharable Content Object Reference Model (SCORM) has become the most popular international standard among the existing ones. In the LOR, a huge
amount of SCORM teaching materials, including associated learning objects, will result in management problems. Su et al. (2005) proposed a management approach called the Level-wise Content Management Scheme (LCMS) to efficiently maintain, search and retrieve learning contents from a SCORM compliant LOR. LCMS includes two phases: the Construction phase and Search phase. In the beginning, the content structure of SCORM teaching materials (Content Package) is first transformed into a tree-like structure called a Content Tree (CT) to represent each piece of teaching material. Based on Content Trees (CTs), the proposed Level-wise Content Clustering Algorithm (LCCAlg) then creates a multistage graph showing relationships among learning objects (LOs), e.g., and a Directed Acyclic Graph (DAG) called the Level-wise Content Clustering Graph (LCCG). This level-wise approach employs the structural information of SCORM-compliant document for retrieval, but the metadata of content packages has not been utilized to increase its precision. Besides, its time and storage complexity have not been addressed in depth.

With the flourishing development of e-Learning, more and more SCORM-compliant teaching materials are developed by institutes and individuals in different sites. In addition, the e-Learning grid is emerging as an infrastructure to enhance traditional e-Learning systems. Therefore, information retrieval schemes supporting SCORM-compliant documents on grid environments are gaining its importance. To minimize the query processing time and content transmission time, a bottom-up approach was proposed to reorganize documents in these sites based on their metadata and to manage these contents in a centralized manner (Shih et al., 2008). An indexing structure named Taxonomic Indexing Trees (TI-trees) as shown in Figure 3 was designed. It is a taxonomic structure with two novel features: 1) reorganizing documents according to the Classification metadata such as queries by classes can be processed efficiently and 2) indexing dispersedly stored documents in a centralized manner which is suitable for common grid middleware. This approach is composed of a Construction phase and a Search phase. In the beginning, a local TI-tree is built from each Learning Object Repository, and then all local TI-trees are merged into a global TI-tree. In the end, a Grid Portal processes queries and presents results with estimated transmission time to users. Experimental results show that the proposed approach can efficiently retrieve SCORM-compliant documents with good scalability. We extend the index scheme to support context-aware learning content retrieval.

![Taxonomic Index Tree for learning content organization](image)

**Figure 3.** A Taxonomic Index Tree for learning content organization

**A Knowledge-based Approach**

The main difficulty for a learner to accurately retrieve relevant teaching material is how to formulate a good query. Usually, the query submitted by a novice user is short and ambiguous, and thus many irrelevant results are retrieved. Accordingly, a knowledge base system (KBS) is very suitable for solving this problem. After the expertise of query formulation is transformed into the knowledge base, the inference engine of the KBS can infer other appropriate keywords to expand the original query and get more relevant results. Furthermore, the knowledge can be reused by others and easily adapted to different scenarios.
To apply a knowledge-based approach to context-aware retrieval, the most important thing is knowledge acquisition. In this section, we present the knowledge transformation model for teachers to use an ontology-driven method to transform their course expertise into query expansion rules. This study assumes that a teacher plans the course outline before s/he teaches a course and lists the concepts to be learned in the class. The Ontology Building Algorithm is designed to assist teachers to transform a course outline into ontology, and then the ontology is transformed into rules by the Ontology_to_Rule Algorithm. In other words, once the course outline is planned, it can be semi-automatically transformed into rules. Therefore, it is a helpful tool for teachers to rapidly generate rules for query expansion.

Problem Formulation

We assume that a context detection module is available, and it can extract users’ context information. Subsequently, several definitions will be introduced, including teaching materials, learning object repository, a query, context and a similarity function.

The symbols in Table 2 are used throughout this paper.

Table 2: The notation used in this paper

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>Content Package</td>
</tr>
<tr>
<td>LOR</td>
<td>Learning Object Repository</td>
</tr>
<tr>
<td>wi</td>
<td>Weighting element i of CP vector</td>
</tr>
<tr>
<td>V</td>
<td>Set of terms in the vocabulary</td>
</tr>
<tr>
<td>Q</td>
<td>Query</td>
</tr>
<tr>
<td>vQ</td>
<td>Vector representing the query</td>
</tr>
<tr>
<td>LCx</td>
<td>The x-coordinate of location context</td>
</tr>
<tr>
<td>LCy</td>
<td>The y-coordinate of location context</td>
</tr>
<tr>
<td>sim()</td>
<td>Similarity function</td>
</tr>
</tbody>
</table>

In the SCORM standard, a Content Package (CP) is defined as a package of learning materials, and a Learning Object Repository (LOR) is a database where the CPs are stored. In this paper, a CP is modeled as a tree to represent the structural information of a CP. To enable content-based retrieval, the well-known Vector Space Model is applied to represent the text content. Besides, SCORM metadata is included in this model of CP as an attribute to utilize the metadata given by authors. Afterward, teaching material, learning content, SCORM-compliant documents and Content Packages are used interchangeably in this paper.

Definition 1 (Content Package). A Content Package (CP) is modeled as a rooted tree where the leaf nodes contain the content and the internal nodes represent the structural information. The content of each level is represented by a vector. In addition, a CP is associated with a set of Metadata.

Example 1

An example CP is modeled as a trinary tree with three levels as shown in Figure 4. The leaf nodes contain the content, and the internal nodes represent the structural information. In addition, a CP is associated with a set of Metadata.

Traditionally, similarity is measured by the Vector Space Model (VSM) in information retrieval domain (Baeza-Yates & Ribeiro-Neto, 1999). In the VSM-based model, a document is represented as a vector. In general, a limited vocabulary of keywords is adopted to denote important words in documents. Each element of the vector corresponds to a keyword of the vocabulary. Therefore, the length of the vector for a document is equal to the size of the vocabulary. The value of each element is a weight denoting the importance of the keyword to the document. There are a number of methods to determine the weights. Among them, TF-IDF (Salton & McGill, 1983; Salton et al., 1975) is the most well-known method to assign weights. The TD-IDF method is based on two findings. First, words frequently used in a document, except stop words, are important keywords to the document. Second, words frequently appearing in many documents are not important for the purpose of differentiation.
As standardization of teaching materials becomes a trend, we model TMs by the SCORM standard, where a Content Package (CP) is defined as a package of learning materials. The content package is represented by a vector of keyword weights. To enable content-based retrieval, the Vector Space Model is applied to represent the text content and to calculate the keyword weights. Also, the Educational category of the LOM metadata is included in this model of CP. Therefore, a TM representation consists of a vector of weights and a category of metadata. Figure 4 illustrates an example of TM representation, where the vector represents weights of 10 keywords and metadata denotes educational attributes. Afterward, teaching materials, SCORM-compliant documents and Content Packages are used interchangeably.

**Definition 2** (Learning Object Repository). A Learning Object Repository is a set of Content Packages located in the same site.

We use weight vector to represent the query, and its formal definition is as follows.

**Definition 3** (Query). A Query is used by a user to specify the TMs s/he wants. Users can express their queries in two forms: keyword-based and metadata-based. A keyword-based query is a vector of keyword weights which mean the concepts about the desired contents. A metadata-based query is a list of (Attribute, Value) pairs, which describe the properties of TMs.

We will now define the notion of similarity between a query and a content package which means the relevance of the content package to the query.

**Definition 4** (Similarity). Let Q be a query with query vector $v_Q$, and TM be a content package. The similarity function is denoted by $\text{sim}(Q, TM)$.

In order to determine the degree of relevance for a query and a teaching material, the similarity function has to be defined. Conventional similarity functions, such as the cosine function, are not suitable for SCORM-compliant teaching materials which are characterized by textual content, metadata and structural information. Hence, a similarity measure $\text{Sim}$ between a query $Q$ and a teaching material $TM$ is proposed by combining a keyword-based similarity and a metadata-based similarity. The keyword similarity $\text{Sim}_{\text{keyword}}$ adopts a cosine function to measure the text similarity between a query and a TM. The metadata similarity $\text{Sim}_{\text{metadata}}$ is defined as the number of matched attributes divided by the number of all attributes. Therefore, the range of these two similarity terms, $\text{Sim}_{\text{keyword}}$ and $\text{Sim}_{\text{metadata}}$ are both in $[0, 1]$. The similarity measure $\text{Sim}$ is defined in (1).

$$\text{Sim}(Q, \text{TM}) = \alpha \times \text{Sim}_{\text{Keyword}}(Q, \text{TM}) + (1 - \alpha) \times \text{Sim}_{\text{Metadata}}(Q, \text{TM})$$

(1)

where the factor $\alpha$, $0 \leq \alpha \leq 1$, is used to control the relative weighting of keyword similarity and metadata similarity.
In terms of ubiquitous learning, it is widely accepted that the contexts include personal and environmental information sensed by the system or retrieved from the databases. Five types of contexts are listed (Hwang et al., 2008; Yang et al., 2007). Among these contexts, “location” may be one of the important information for describing a learner’s status (Lonsdale et al., 2003). In this research, we focus on the location context sensed by the system, which is represented by (x, y)-coordinates. Location-awareness means the system is aware of the learner’s location by means of technologies such as the Global Position Services (GPS). We assume that context information of location can be acquired by extensively deployed sensors and built-in maps.

**Definition 5** Location Context. The Location Context is represented by a two-dimensional coordinate, \((LC_x, LC_y)\), where \(LC_x\) is the x-coordinate, and \(LC_y\) is the y-coordinate. These coordinates correspond to a map of the campus.

Based on the definitions above, the Context-aware Learning Content Retrieval Problem (CALCRP) can be defined as follows.

**Definition 6** Context-aware Learning Content Retrieval Problem (CALCRP). Given a query and context information, this problem is to retrieve relevant learning contents from a repository ranking by a similarity function. The goal is to improve precision and recall of retrieval.

**Ontology Building**

Ontology building has been considered as a craft rather than an engineering activity. Traditionally, the process of ontology building requires the participation of domain experts and knowledge engineers. Although a number of automatic technologies of ontology construction have been proposed, it is still not easy for teachers and domain experts to build up ontology. Therefore, the ontology building algorithm is proposed for teachers to easily derive ontology from course outline. In this algorithm, an “expert” means an educator who is also good at knowledge engineering.

Before describing the process of ontology building, we give a general definition of ontology.

**Definition 7** (Ontology). Ontology is a conceptualization of a domain, which is defined as a quadruple \(O= (C, A, R, X)\), where
- \(C\) is a set of concepts;
- \(A\) is a collection of attributes sets, one for each concept;
- \(R\) is a set of relations on \(C \times C\);
- \(X\) is a set of axioms.

**Example 3**
A Campus_Plant_Course Ontology \(O_{CPC}\) = \((C, A, R, X)\) is an ontology where its components are endowed as follows.
- \(C = \{\text{“Plant,” “Structure,” “Fern,” …}\}\)
- \(A = \{\text{Keyword, Type, Location, Level}\}\)
- \(R = \{\text{“is_a,” “related_to”}\}\)
- \(X = \{\text{IF is_a(“A”, “B”) and is_a(“B”, “C”) THEN is_a(“A”, “C”)}\}\)

In this study, the ontology is derived from a pre-defined course outline which reflects the content of the course to be taught by the teacher. The course outline is usually organized by the teacher before the class begins. Moreover, a course outline is defined as a two-level structure, chapters and sections.

**Definition 8** (Course Outline). A Course Outline is a two-level tree-like representation of the table of content for a course. A course outline consists of a limited number of Chapters, which consists of a limited number of Sections.

**Example 4**
A Campus_Plant Course Outline \(CO_{CP}\) can be represented as follows.

Course Name: Plants in the Campus
Chapter 1. Introduction to Plants
    Section 1.1. What is plants?
    Section 1.2. Classification of Plants

Chapter 2. Structures of Plants
    Section 2.1. Flowers
    Section 2.2. Leaves
    Section 2.3. Fruits

Chapter 3. Growth of Plants
    Section 3.1. Budding
    Section 3.2. The Growing Process

Chapter 4. Identification of Plants in the Campus

After the course outline is determined, the teacher can follow the steps of the Ontology Building Algorithm to derive ontology from the course outline. This is a special-purpose algorithm which is designed for constructing the ontology of a course about plants in a campus. Teachers who teach this kind of courses can follow this algorithm to generate ontology.

Subroutine Topic_Concept_Extraction (Sub_TCE) extracts the topic concept according to the name and content of a chapter by calling a subroutine which is designed based on heuristics. First, the name of the chapter is scanned to extract a keyword representing the topic concept of the chapter. If the given name of a chapter does not represent or illuminate any related concept, the text content of this chapter is represented by the Vector Space Model, and the TF-IDF (Term Frequency-Inversion Document Frequency) scheme is used to find the keyword with the largest weight. Then, this keyword is assigned as the topic concept of the chapter. Likewise, the course concept can be extracted in a similar manner.

Subroutine Topic_Concept_Extraction (Sub_TCE)

Symbols Definition:
Chap_Name: the name of the chapter
Chap_Content: the Vector-Space-Model representation of the chapter
Topic_Concept: the extracted topic concept

Input: Chap_Name, Chap_Content
Output: Topic_Concept

Step 1: Extract the related Topic concept from the name of the chapter.
   Step 1.1: Remove the stop words.
   Step 1.2: Extract the keywords from the name of the chapter in stem forms.
   Step 1.3: Remove the Course concept from the keywords.
   Step 1.4: If only one keyword remains, then assign it as Topic_Concept, and return.

Step 2: Extract the related Topic concept from the content of the chapter.
   Step 2.1: Use the TF-IDF scheme to find the keyword with the largest weight.
   Step 2.2: Assign it as Topic_Concept, and return.

The Ontology Building Algorithm is listed as follows, which can be executed by computers after appropriate implementation.

Campus_Plant_Course (CPC) Ontology Building Algorithm

Symbols Definition:
Course_Outline: a two-level course outline as defined in Definition 8
CPC_Ont: an ontology for a course about campus plants, as defined in Definition 7

Input: Course_Outline
Output: CPC_Ont

Step 1: Build the skeleton CPC ontology.
   Step 1.1: Extract the Course concept from the name of the course.
Step 1.2: For each chapter, call Sub_TCE to extract the Topic concept.
Step 1.3: For each Topic concept, create a “related_to” relation to the Course concept.

**Step 2:** Develop concept hierarchy for the Course concept with “is_a” relations.
- Step 2.1: Create one Category concept for each kind of plants appearing in the course content.
- Step 2.2: For each Category concept, create an “is_a” relation to the Course concept.

**Step 3:** Develop concept hierarchies for Topic concepts.
- Step 3.1: For each Topic concept, extract Sub-topic concepts from the names of sections.
- Step 3.2: For each Sub-topic concept, create a “sub_topic” relation to the Topic concept.

**Step 4:** Experts verify the ontology.

Associating concepts with instances, also called Annotation, is an important task for retrieving learning content based on the ontology. However, it is really heavy work for teachers to annotate all plants in a campus. Therefore, we have reduced the workload of a teacher to a reasonable degree. For example, the teacher only has to identify instances of plants appearing in the course content. In other words, all the teacher needs to do is to identify one Instance concept in the campus for each category in the ontology. In the future work, we plan to propose a collaborative approach to annotating plants in a campus-wide manner. By an appropriate distribution of workload, a whole ontology of plants in the campus can be built up and annotated.

The process of ontology building is shown in Figure 5. Two types of relationships in the Campus_Plant_Course ontology (CPC Ontology) are described as follows. (1) “is_a” is a generalization relationship, which could be used to describe the concept taxonomies in the class hierarchy. For example, either a wooden plant or a fern plant is a kind of plants. (2) “related_to” denotes that there exists a “related_to” relationship between two concepts. For example, we could use the “related_to” relationship to denote that the “plant” subject is related to the “structure” topic.

![Figure 5. Process of ontology building](image)

This section describes the process of constructing the knowledge base. Throughout this paper, we take the campus plants for example. Just like the concept of object orientation, we could view all of the entities in the campus as concepts and it is natural for us to model the campus using concept hierarchies. For example, a “woody plant” is a concept, and it contains attributes or slots: Leaf Shape, Leaf Color, Leaf Size, etc. Furthermore, people tend to group the knowledge and build up structural information when they learn new concepts. The grouped knowledge could be viewed as a bigger concept as well. For example, both woody plants and fern plants are typical types of plants. Hence, the “woody plant” concept inherits the “plant” concept, and there exists a relationship between them. In essence, ontological representation is suitable for communications and natural for human thinking; meanwhile, rule-based representation is powerful for machine to manipulate the concepts. As described above, ontology could be
used to model the concept hierarchy and relationships between concepts. However, it is not easy to model the behavior of concepts using ontology only. When the problem domain is described clearly and well modeled, it is much easier to build up a rule-based expert system because many tools (called expert system shells) can offer assistances. Hence, in practice, rule-based representation is more suitable for building applications. On the other hand, since most real-world applications need complex rules to model, the meaning captured into ontology for the problem domain becomes very helpful for rule extractions when building complex systems.

System Overview

We have proposed a context-aware learning content retrieval system based on the knowledge transformation model. As shown in Figure 6, the overall system consists of four components:

- User Interface: query input and context detection
- Query Expansion: expanding a query by rule inference
- Content Retrieval: searching and results ranking
- Knowledge transformation: ontology building and rules generation

The flow of the system can be summarized as follows. First, the query and context information are transferred to the Query Expansion component to generate an expanded query. Next, the expanded query is sent to the Content Retrieval component for query processing and retrieving relevant content. Finally, the retrieved results are returned to the user. The search process is carried out by the Query Processor component, which receives expanded queries, processes the queries and presents results to the users.

In this work, we assume that the instructional strategies and corresponding retrieval strategies have been defined in the knowledge base by experts. Strategies of query expansion include:

- Specialization: Giving keywords belonging to subordinate concepts
- Generalization: Giving keywords belonging to super-ordinate concepts
- Association: Giving keywords belonging to related topics

Knowledge is represented by rules describing actions of query expansion. After the generation of ontology, we derive rules for query expansion using an ontology-driven method. Rules are automatically transformed from the ontology and extracted from two kinds of relations:

- “is_a” relation: For generalization/specialization strategies
  For example, IF (Strategy="Specialization") and (term="Fern") THEN expand("Fern 1").
• “related_to” relation: For association strategies
  For example, IF (Strategy="Association") and (subject="Plant") THEN expand("Structure").

The algorithm is listed as follows.

Ontology_to_Rule Algorithm
Input:
  CPC ontology
Output:
  Rules for query expansion

Step 1: Extract rules by “is_a” relations.
1.1: for each “A ‘is_a’ B” relation, generate:
  IF (Strategy='Generalization') and (term="A") THEN expand("B")
1.2: for each “A ‘is_a’ B” relation, generate:
  IF (Strategy='Specialization') and (term="B") THEN expand("A")

Step 2: Extract rules by “related_to” relations.
2.1: for each “A ‘rel’ B” relation, generate four rules:
  IF (Strategy='Association') and (Subject="A") THEN expand("B")
  IF (Strategy='Association') and (Subject="B") THEN expand("A")
  IF (Strategy='Association') and (Topic="A") THEN expand("B")
  IF (Strategy='Association') and (Topic="B") THEN expand("A")

Step 3: Verify the generated rules by domain experts and ask the knowledge engineers to modify the rules if needed.

To support the pre-defined instructional strategy and enable context-awareness, we propose to divide the process of query expansion into two phases:

• Phase 1: Candidate keywords generation. Based on the retrieval strategy derived from the instructional strategy, candidate keywords are recommended. This phase extends the instructional strategy.
• Phase 2: Context-aware filtering. This phase focuses on filtering the candidate keywords to realize context-awareness. Among a variety of methods, we adopt a distance-based method to determine relevance of keywords.

The main advantage of dividing the process of query expansion is the separation of pedagogic design and technical implementation. While the first phase generates keywords based on instructional consideration, the second phase is related to technical factors. Various technologies can be applied if necessary. Actions of query expansion can be based on various ideas. The point is how to determine the correlation of two keywords. Conventional methods of calculating keyword correlation include: thesaurus-based, co-occurrence in the corpus and top-ranked in the returned set. Our idea is based on geographical proximity. The expanded keyword is mainly related to the learner’s location, instead of the original query.

The distance of two entities is defined as follows.

\[
\text{Dist}(C, L) = \sqrt{(x_C - x_L)^2 + (y_C - y_L)^2}
\]  

(2)

Dist(C, L) means the geographical distance between C and L, where
• C: a concept;
• L: a learner;
(x_C, y_C), (x_L, y_L): coordinates of C and L.

As shown in Figure 7 and 8, for specializing “Fern,” there are three choices: Fern 1, Fern 2 and Fern 3. We choose the nearest one to the learner.
The User Interface receives users' queries and context information extracted by the Context Detection Module. Another task of the User Interface component is initialization of facts for inference. For example, the fact of the strategy defined by teachers is initialized and loaded into the knowledge base for inference. Another important task is to initialize the campus map for default reasoning. The campus map records coordinates of primary plants, buildings and other entities. Although we assume that a lot of sensors have been installed in the u-learning environment, it is probable that the user walks by an area without a sensor. In this situation, coordinates in the campus map can serve as default context information to prevent the inference process from being failed.

An Illustrative Scenario

This section presents a scenario of a context-aware ubiquitous learning environment, where some communication and context-sensing devices have been installed to enable context-aware retrieval of teaching materials. For an “identification of plants” class of an elementary school, the teacher sets up five learning corners in the campus. The students in class are divided into five groups, and the teacher arranges an on-line tour guidance for these students. When a student equipped with a PDA goes to the first corner, the system interacts with the student:

- System: Can you identify the plant in front of you?
- Student: Yes.
- System: Can you identify the name of this plant?
- Student: Yes.
- System: What is the name of this plant?
- Student: Fern.
- System: Can you describe the characteristics of a fern?
After the student submits the query, the system conducts the context-aware retrieval and then presents relevant content to the student.

In this scenario, the system retrieves relevant teaching material according to the location context of the student. For example, assuming the fern plant in the first corner is some kind of fern. We may call it Fern 1. We assume that the contents in the repository include teaching materials about Fern 1, Fern 2 and Fern 3. The system will deliver content about Fern 1 to the student in accordance with the context information. The content describing the plant seen by the student will strengthen the impression of the learner and therefore enhance the learning effect.

Also, we think the similarity can improve the learning performance. The query reflects the things the student wants to learn. The more questions of the learner can be solved when he/she is provided with more relevant content, the more the learning effect can be increased. Context-aware retrieval can provide students with relevant teaching materials and shorten the learning process.

The interactive query mechanism is a common process proposed in ubiquitous learning activities. However, the design of a meaningful question base has not been widely discussed in the literature. In our opinion, the system can ask questions of the learner according to a pre-defined scenario generated from the course objectives and the teaching material. For example, if the course objective is to identify the name of some plant mentioned in the teaching material, the scenario will contain such a question as “What is the name of the plant?” In the future work, we plan to make the course objectives converted automatically and fit the teaching material into the query scenarios which can be refined by teachers if necessary.

**Method**

Our proposed approach has been implemented and the prototype was provided to several teachers of an elementary school for evaluation. The corpus is composed of SCORM-compliant teaching materials adopted from those of repositories built by Taiwan Ministry of Education (http://nature.edu.tw). The experiments investigate the accuracy of searching. Also, some questionnaires involved in such issues are used to respond to the user satisfaction.

**Prototype**

To evaluate the proposed approach, we have implemented a web-based prototype. As shown in Figure 9, users can submit queries in this web page. Then, the original query is expanded by the knowledge-based system. After that, the required contents are retrieved from repositories and delivered to the user. The retrieved content packages are ranked by their similarities to the query. All programs are implemented in the Java language.

In this paper, we use DRAMA (Lin et al., 2003) as an expert system shell because of its client–server architecture and the object-oriented knowledge base structure. The purpose of the DRAMA’s server is to load, manage, and use the knowledge bases according to the knowledge service that users need. DRAMA’s server contains many different rulebases and provides different APIs for the application servers to connect. The application server employs DRAMA’s APIs to provide user-friendly web pages for users to use expert systems. Based on the client-server architecture, it becomes very easy for us to develop a KBS for supporting context-aware u-learning. The Lucene search engine (http://lucene.apache.org) was adopted to perform basic keyword indexing, search and retrieval in the prototype. This IR engine was open-source software developed by the Apache Lucene project. Lucene is characterized by the ease to enable applications to index and search documents.
There are nearly one hundred Content Packages in the LOR, which are retrieved and adopted from existing repositories on the Internet like http://learning.edu.tw/mainpage.php. Currently, the LOR is only available to elementary school teachers who participate in this evaluation. However, in the near future, we plan to place the prototype and the LOR on the web for public access and large-scale evaluation.

The concept of “relevance” of a TM to a query has to be defined before the retrieval method can be evaluated. It is somewhat subjective, depending on the users, to judge whether a document is relevant to a query or not, especially in a dynamic changing environment. For example, a document about fern is usually thought of lowly relevant as a query of “tree.” However, while the user is standing in front a fern plant, this document will be highly relevant to a query. In this experiment, we try to make an objective, reliable and fair measure for the relevance. First, we define that a document and a query is relevant if their similarity value calculated by (1) exceeds the threshold value assigned in advance in the system. Next, for 5 places in the campus, we manually generate a query for each of them. Then, the documents in the repository are manually judged its relevance to the 5 queries by teachers and experts.

In this experiment, we use two well-known metrics of information retrieval, precision and recall, to measure performance of the proposed approach. We define precision and recall as follows.

\[
Precision = \frac{R_{ret}}{Ret} \\
Recall = \frac{R_{ret}}{R_{LOR}}
\]

where

- \(R_{ret}\) is the number of relevant documents in the retrieved documents;
- \(Ret\) is the number of retrieved documents;
- \(R_{LOR}\) is the number of all relevant documents in all repositories.

To optimally set up the parameters of the similarity function, \(\alpha\) is not easy. In this experiment, we set \(\alpha = 0.5\). This setting means equal weighting for keyword similarity and metadata similarity. Other settings of the parameter will be considered in future work.
Study Design

In this section, three experiments/surveys are conducted to address respectively as follows:

- the performance of knowledge-based query expansion;
- the learning performance of using the proposed system; and
- the satisfaction survey of using context-aware retrieval.

The purpose of this first experiment is to evaluate the performance of knowledge-based query expansion with respect to the precision and recall of context-aware content retrieval. The participants are 15 fourth-grade students invited from an elementary school in Nantou of Taiwan. After one-week usage training, they use this system to retrieve relevant teaching material. At each learning corner in the campus, every student submits a query. Then, the original query is transformed by the system into a specialized query and a generalized query respectively. Next, the three queries are used to retrieve teaching materials, and the precision and recall values are calculated. Figure 10 illustrates the average precision and recall values for the original query, the specialized query, and the generalized query. The results show that the expanded queries perform better than the original one. The main reason may be that the original queries are usually short and ambiguous, and therefore they are insufficient to represent the intention of users. In addition, we found that generalization can improve recall, and specialization can improve precision. This is consistent with the cognition of precision and recall.

The second experiment was conducted to evaluate the effects of the location-aware learning content retrieval system. This experiment specifically investigated the following question:

“Can location-aware learning content retrieval contribute to students’ learning?”

This experiment was conducted at an elementary school, and the subjects were 20 fourth-grade students from the same class taught by the same teacher. The experimental course is named “Introduction to Campus Plants,” which aims to help students learn knowledge about plants in the campus. Before the experiment conducted, students were trained to be familiar with the retrieval tool. The training consists of two phases. First, all students learned in a conventional classroom and used the learning content retrieval system without the location-awareness feature. Then, they were educated how to access the retrieval tool through a PDA. After training, the 20 students were separated into two groups: A (control group) and B (experimental group), each of which contained 10 students. The students in Group-A had access to learning content retrieval system without the location-aware feature while those in Group-B had access to the proposed location-aware learning content retrieval system. However, the location-awareness feature is transparent to students. That is, users do not know that the tool can retrieve learning content according to their location context. Before the course began, the 20 students were given a pre-test to ensure that they had the equivalent performance for the course. After the course finished, a post-test was given to respond to the comparison of the learning achievements. The contents of the two tests cover those learning materials related to the “Introduction to Campus Plants” course in the Learning Object Repository. Common threats to internal validity include: Confounding, Selection (bias), History, Maturation, Repeated testing, Instrument change, Regression toward the mean, Mortality/differential attrition, Selection-maturation interaction and Experimenter bias. In this experiment, we are careful to avoid these threats. The statistical results from applying SPSS to analyze the tests are presented in the next section.

The items of the pre-test include 25 single-choice questions, and sample items are shown in Appendix A. In this study, the items were presented in Chinese when undertaking this study, and the test items, shown in the Appendix, were translated by the authors. We went through five stages during the process of test construction: item construction, preliminary test, item analysis, item selection and validity analysis. First, a two-way specification table was constructed according to learning content and Bloom’s Taxonomy of Educational Objectives (Bloom, 1956) as shown in Table 3. The content of the test covers the following levels of expertise: knowledge, comprehension, application, analysis and synthesis. The draft of the pre-test items includes 30 single-choice questions. 30 fourth-grade students who did not participate in the pre-test and the post-test are invited to take the preliminary test. According to the results of the preliminary test, we eliminate 5 questions which have low difficulty and discrimination. The two-way specification table shows the learning units and the educational objectives. After the initial construction of the pre-test items, two teachers in the fields of Nature Science were invited to make comments on it for face validity.
Table 3: The two-way specification table of the pre-test

<table>
<thead>
<tr>
<th>Learning Unit</th>
<th>Level</th>
<th>Knowledge</th>
<th>Comprehension</th>
<th>Application, Synthesis</th>
<th>Analysis and Synthesis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functions of root/stem/leaf</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>Functions of flower/fruit/seed</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>7</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>5</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The items of the post-test were constructed in a way similar to the pre-test, as shown in Appendix B. The two-way specification table of the post-test is shown in Table 4.

Table 4: The two-way specification table of the post-test

<table>
<thead>
<tr>
<th>Learning Unit</th>
<th>Level</th>
<th>Knowledge</th>
<th>Comprehension</th>
<th>Application, Synthesis</th>
<th>Analysis and Synthesis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutritional organs of plants</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Reproduction of plants</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>11</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>5</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The two variables that may contribute to the difference between the two groups are carefully controlled in the experiment.

- The learning environment.
The students of the two groups used the same retrieval system in the same outdoor campus. The only difference is that the context-awareness feature of the retrieval system was disabled for the control group.
- Teachers and learning activities.
Students were taught by the same teacher during the experiment. Additionally, the instruction or facilitation received by students was not different.

Next, a survey was conducted with respect to the 15 students in the first experiment. The questionnaire contains questions of four categories, each of which has five questions. The purpose is to contemplate their comments on the retrieved contents and learning effect. Besides, a five-point Likert scale with anchors ranging from strongly disagree (1) to strongly agree (5) is used for this survey.

**Experimental Results**

In this section, the results of three experiments/surveys are reported.

The purpose of this first experiment is to evaluate the performance of knowledge-based query expansion with respect to the precision and recall of context-aware content retrieval. At first glance, the results shown in Figure 10 seem to conflict with conventional information retrieval practice which indicates the trend of decreasing precision along with the rise of recall. In fact, there is no conflict. To generate a typical precision-recall plotting for a given query, the set of retrieved documents are listed. Next, the precision and recall are calculated accumulatively from the first document to the last document. Finally, these pairs of precision and recall values are plotted in a 2-dimensional coordinate figure with the precision against the recall. In this kind of figures, the following trend usually holds, as the recall rises, the precision decreases. However, Figure 10 is not obtained in this manner. Given a collection and a query, the precision and recall values are calculated by (3) and (4). That is, while we focus on one query, conventional experiments illustrate the results of multiple queries. Therefore, it is possible for the expanded query to outperform the original query in both precision and recall.
The second experiment was conducted to evaluate the effects of the location-aware learning content retrieval system.

- **Pre-test**
  The aim of the pre-test is to ensure that both groups of students had the equivalent performance for the course. Table 5 presents the t-test results of the pre-test. The mean scores for the pre-test reveal that Group A performed as well as Group B. That is, |t| = 1.37 < t_{\alpha}(9) = 1.833, which implies that the performance of Groups A and B in the pre-test does not differ significantly. Therefore, we can conclude that Group A performed as well as Group B in the pre-test conducted before performing the experiment.

  \[
  \begin{array}{cccc}
  \text{A (Control Group)} & 10 & 52.30 & 6.67 \\
  \text{B (Experimental Group)} & 10 & 47.70 & 7.06 \\
  \text{A - B} & 10 & 4.60 & 10.64 & 1.37 \\
  \end{array}
  \]

- **Post-test**
  The post-test aims to compare the learning achievements of the two groups of students after taking the course with different retrieval tools. Table 6 shows the t-test results for the post-test. From the mean value of the post-test, Group B performed better than Group A. |t| = 1.98 > t_{\alpha}(9) = 1.833, which implies a significant difference between the performance of Groups B and A in the post-test. Therefore, we can conclude that Group B achieved a significant improvement compared to Group A after using the location-aware learning content retrieval system.

  \[
  \begin{array}{cccc}
  \text{A (Control Group)} & 10 & 83.40 & 16.08 \\
  \text{B (Experimental Group)} & 10 & 91.70 & 5.83 \\
  \text{A - B} & 10 & -8.30 & 13.23 & -1.98 \\
  \end{array}
  \]

Next, a survey was conducted with respect to the 15 students in the first experiment. The mean value and standard deviation (SD) are calculated for each category, as shown in Table 7.

\[
\begin{array}{ccc}
\text{Category No.} & \text{Questions} & \text{Mean} & \text{SD} \\
1 & \text{Satisfaction of the user interface of the system} & 2.9 & 1.06 \\
2 & \text{Satisfaction of the retrieved contents and learning effect} & 4.3 & 0.89 \\
3 & \text{Willingness to use this system for learning} & 3.7 & 1.03 \\
\end{array}
\]
For questions of Categories 1 and 3, the deviation of user satisfaction is slightly larger than other categories. The reason may be that the participants are not all familiar with the usage of the system and some English interfaces. Some participants comment that they are not used to studying on computers. However, some participants appreciate this idea and like to retrieve relevant TMs.

The results of Categories 2 show that most participants are satisfied with the expanded queries and the retrieved contents. Most students agree that the expanded queries can enhance their original queries, and the retrieved contents are helpful to their learning. In summary, the system can help students efficiently find relevant teaching materials for learning. However, the user interface has to be improved to attract more users.

Discussion

Based on the assumptions mentioned in the Introduction Section, the experimental results can be interpreted as follows. First, the location-aware retrieval tool can provide students with “right” content about their nearby plants, insects, etc. We think that the right content can promptly solve students’ questions. Furthermore, the knowledge related to such nearby objects could be more impressive and interesting to increase users’ learning performance. Based on the assumptions, the experimental results can be interpreted as follows:

- The location-aware retrieval tool can provide students with “right” content about their nearby plants, insects, etc. We think that the right content can promptly solve students’ questions and that the knowledge related to nearby learning objects could be more impressive and interesting to improve the learning performance.
- The integration of physical situation into instruction can enhance students’ learning performance. This viewpoint can also be found in the model of “Situated Learning” proposed by Lave & Wenger (1991). Traditionally, learning has been viewed as transmission of knowledge from teachers to students. However, the Situated Learning Theory argues that learning should be situated in a specific contextual, social and physical environment.
- The “seamless” feature of ubiquitous learning makes the learning process more convenient than the conventional learning. For example, in an ideal u-learning environment, a student with an RFID-enabled device can experience “seamless learning at right time and right place.”

The findings of experimental results are interpreted in terms of related literatures. First, we address the application of query expansion to context-aware retrieval. Query expansion has been investigated to improve recall of information retrieval and disambiguate the meaning of queries. A large number of researches have been devoted to this topic (Bhogal et al., 2007), but query expansion has not been widely applied to the e-learning domain and mentioned in the ubiquitous learning. Shih et al. (2008) indicated that efficient retrieval of teaching material could facilitate the learning process and then enhance the learning effects. In this study, the result of the first experiment implies that the proposed query expansion can improve the performance of context-aware retrieval, which can support the ubiquitous learning scenario.

Currently, researches of context-aware ubiquitous learning focus on the acquisition and modeling of context information (Oppermann & Specht, 1999; Yang, 2006; Yin et al., 2005), such as location, temperature, and humidity etc. In this study, we find that the context-aware retrieval tool is helpful to students’ learning. They comment that the system can retrieve context-related contents, which saves their time to find relevant references.

The knowledge engineering like ontology building has been thought of as tough work which can only be dealt with by domain experts and knowledge engineers. Therefore, researches of ontology building focus on automatic or semi-automatic approaches (Liu et al., 2004; Tho et al., 2006) to alleviate the burden of the builders. This study proposes a teacher-guided approach to build simple ontology for educational usage. The result of survey shows that it is feasible for teachers to provide their expertise and help the system generate a simple ontology based on a pre-defined course outline.

In addition, the contribution of this paper about pedagogical feasibility and keyword association is clarified. First, to derive the important components of the proposed approach, the ontology and the knowledge base, an automatic approach is adopted to alleviate the burden of teachers and domain experts. For the proposed setting, teachers guide the automatic construction of the ontology by providing a course outline and instances of concepts existing in the campus. Next, domain experts verify the ontology generated by the proposed algorithm. We do not intend to require
teachers to manually build up the rules and ontology. Instead, teachers provide their knowledge about course outline and campus context, which is not difficult, and the system will transform them into ontology and rules. In this way, the proposed approach will be pedagogically feasible. Second, the proposed approach is distinct from existing searching engines. In particular, such as Google, we can not adaptively find results through users’ context. Although collaborative filtering techniques have succeeded in suggesting contents of keyword association mining from users’ query logs, the context-awareness has not been integrated into this technology.

Conclusion

With the flourishing development of e-learning, more and more learning object repositories are constructed and connected to share the content. Efficient content retrieval schemes, such as context-aware retrieval, can reduce the response time and thus attract more users to utilize the e-learning systems. The existing methods of information retrieval can not rapidly and accurately satisfy the request of retrieving desired learning contents in context-aware ubiquitous learning environments. The proposed approach attains rapidness and precision by a knowledge-based approach with which the desired learning contents can be efficiently retrieved and then advance the sharing and reusing of learning contents. For example, one application of this technology is to teach material design. To support individualized and adaptive learning, teachers are encouraged to develop various teaching materials by means of different requirements. However, traditional methodologies for designing teaching materials are time-consuming. To speed up the development process of teaching materials, teachers can reuse existing contents to rapidly generate context-aware contents.

We have designed a knowledge-based approach to query expansion for context-aware learning content retrieval. In addition, a prototype was built to implement the model. Experimental results show the proposed approach can enhance original queries to improve the precision. We have also shown that explicit context and content ontology can be used during knowledge transformation for dealing with the complexity of context-aware knowledge management, especially for guiding the construction of the knowledge base. Our main contributions are: (1) to design and implement a framework of context-aware learning content retrieval for supporting context-aware u-learning; (2) to propose an ontology-driven model for eliciting rules from a previously built ontology and constructing the knowledge base. According to the experimental results, the paradigm of using ontology to build up a context-aware learning content retrieval system works well and effective. This system will benefit from the sharing and reusing of built knowledge, the reduction of people’s time to learn knowledge management, the ease of context-aware course design and the improvement on the precision of learning content retrieval.

The main findings from the experimental results can be summarized as follows: First, the proposed knowledge-based query expansion can improve the performance of context-aware retrieval. Second, using context-aware retrieval can efficiently retrieve relevant teaching materials which are helpful to learners. The results of the experiment in which 20 elementary school students participated reveal that some students, compared with the others in the control group, using the location-aware learning content retrieval system made significant progress in learning. Therefore, we conclude that the proposed approach can help students improve their learning performance.

On the issue of knowledge maintainability, the proposed approach is based on an object-oriented rule model (NORM), which facilitates knowledge maintenance and reuse. When the scenario or contextual environment is changed, the designers can update the original ontology and the new rules will be automatically generated. It is supposed that the same approach could be adaptively modified to other contextual scenarios for knowledge base construction. Moreover, teaching material design is a sustainable and evolving task that the contextual knowledge of the learning environment might need to be updated from time to time.

Location-aware learning content retrieval can be applied to various ubiquitous learning activities as shown in Table 1, such as museum guide and laboratory assistance etc. In these applications, the readers can easily apply the system to their teaching activities. For example, the learner can retrieve relevant content about the artist’s work in front of her/him during museum guidance.

Contexts of learning status and learning requirements are very important clues for retrieving learning materials. Hence, we plan to consider these contexts acquired by Learning Management Systems for context-aware learning content retrieval. For example, a Learning Portfolio Mining (LPM) system was developed to extract learning features
With this context information, the proposed knowledge-based approach can be easily extended to consider the educational contexts by substituting learner-related knowledge such as learning status and learning requirements and then to adaptively provide appropriate content for learners. Currently, this work focuses on the location context. However, this approach can be extended to consider more types of context information.

In the experiment, the only difference between the control and experimental group is the “location-aware feature.” Namely, students of the two groups use the proposed knowledge-based tool with the same retrieval interface. While the experimental group is implicitly supported with the location-aware feature, the control group can only retrieve results which are not refined from the location context. The proposed knowledge-based approach is flexible and can be extended to process other kinds of context information such as personal profile, time, location and temperature etc. This work focuses on the location context. Therefore, the experiment was carefully set to avoid the interference of other features. Furthermore, the experiment addresses the impact of location-aware content retrieval on students’ learning. We believe that other context features could also improve students’ learning, which will be investigated in our future work.

In addition, our future work will investigate knowledge acquisition for mapping instructional strategies to retrieval strategies. Next, we plan to extend this approach to model personalized learning content retrieval and hence to facilitate an adaptive learning environment. Also, the proposed approach will be extended to model other types of context information, such as time, activities, and peers etc. In this paper, we focused on the management of learning contents which have three types of attributes: textual contents, metadata and structural information. SCORM is one of the standards which satisfy the requirements. In fact, improving the proposed method to consider the learning design of contents such as IMS-Learning Design is another interesting issue of our future work.

Acknowledgements

A preliminary version of this paper was presented at Technology Enhanced Learning Conference 2007. This research was partially supported by National Science Council of Republic of China under the number of NSC95-2520-S009-007-MY3 and NSC95-2520-S009-008-MY3.

References


Appendix A. Selected Sample Items of the Pre-test

1. (   ) Which is the main function of roots? (1) Water absorbing (2) Water transpiration (3) Nutrients making (4) Nutrients transportation.
2. (   ) Which function do potato and taros have in common? (1) Climbing (2) Breathing (3) Nutrients storage (4) Insect catching.
3. (   ) Which sentence is false? (1) Some plants may have metamorphosed leaves, which have special functions. (2) Stems of most plants are like long tubes for water transportation. (3) Roots of most plants are thin and branching for absorbing water. (4) Leaves of most plants are flat for absorbing water.
4. (   ) Which is the main function of seeds? (1) Photosynthesis (2) Nutrients transportation (3) Growing new plants (4) Water absorbing.
5. (   ) What is the purpose of blooming and fruiting? (1) Food for humans (2) Food for animals (3) Reproduction (4) Nutrients consumption.

Appendix B. Selected Sample Items of the Post-test

1. (   ) Most plants have vascular bundles and belong to (1) Liliopsida (2) Magnoliopsida (3) Non-Vascular plants (4) Vascular plants.
2. (   ) Which sentence is true? (1) The leaves of the cactus are thick and large. (2) The leaves of the cactus store a lot of water. (3) The cactus can grow well in dark and wet area. (4) The needle-like leaves of the cactus is to reduce the consumption of water.
3. (   ) Which sentence about potato and sweet potato is true? (1) Both are roots. (2) Both are stems. (3) The part we eat can store nutrients. (4) The part we eat can produce nutrients.
5. (   ) Which is not the feature of ferns? (1) Non-blooming (2) Non-fruiting (3) Reproduction using spores (4) Found at dry and warm places.
Ubiquitous Performance-support System as Mindtool: A Case Study of Instructional Decision Making and Learning Assistant

Hsinyi Peng and Po-Ya Chuang
Institute of Education, National Chiao Tung University, Hsinchu 300, Taiwan // hpeng@mail.nctu.edu.tw // poya@dato.idv.tw

Gwo-Jen Hwang*, Hui-Chun Chu, Ting-Ting Wu and Shu-Xian Huang
Department of Information and Learning Technology, National University of Tainan, Tainan city 70005, Taiwan // Tel: 886-915396558//Fax: 886-3017001 // gghwang@mail.nutn.edu.tw // carolchu@mail.nutn.edu.tw // danytingting@gmail.com // rumb-73716@yahoo.com.tw

*Corresponding author

ABSTRACT

Researchers have conducted various studies on applying wireless communication and ubiquitous computing technologies to education, so that the technologies can provide learners and educators with more active and adaptive support. This study proposes a Ubiquitous Performance-support System (UPSS) that can facilitate the seamless use of powerful new technologies in the school setting. In order to help the readers visualize these novel technologies in practice, we present one case study of a butterfly-ecology training course facilitated by the UPSS. The aim behind the case study is to inform the design and the development of context-aware ubiquitous computing system and its learning materials. The research inquiry centers around three themes: (1) the critical features to the data-driven decision making of teachers, (2) the perceptions of teachers and students to the UPSS, and (3) implementation issues. The results of the two rounds of formative evaluation indicate positive effects of the UPSS regarding motivation, interactivity, and effectiveness. In addition, teachers’ attitudes and teachers’ pedagogical approaches toward UPSS use are two key factors in the successful implementation of teaching with such innovative technology. This study can be a useful reference for those who are interested in conducting studies applying context-aware ubiquitous computing to educational contexts. Finally, this study presents suggestions and implications for future research and system development.

Keywords
Ubiquitous learning, Context awareness, Data-driven decision making, Performance-support system, Ubiquitous performance-support system

Introduction

In 2001, the United States government passed the No Child Left Behind Act (NCLB), which requires states to develop annual assessments of school and student progress that, as accumulated data, are to help educators improve the learning of all students. Similar to NCLB, policy that the government in Taiwan has established requires administrators and teachers to use data to help improve the quality of education on the island (Hwang, 2003). The educators are confronting complex sources of data from which the educators must make informed instructional decisions (Cagiltay, 2006; Hwang et al., 2008). These new expectations have placed heavy responsibilities on educators, who now are trying not so much to modify as to re-conceptualize educational decision making and who now, also, are simultaneously extending data to these newly re-conceptualized areas, ranging from resource allocation to instructional practices. As a result, educators need adequate performance support to facilitate this transition so that they can become successful in this new working environment.

Recent digitalization around the world has been proceeding toward wireless sensor networks, which embed computation and communication components into the environment. These devices detect certain aspects of the contexts of our daily lives, and provide personal support accordingly. Such technology has been called ubiquitous computing (u-computing). In the meanwhile, advances in these new technologies have led to a new research issue in education, that is, the issue of developing a novel learning environment so that students can learn in any place at any time. Moreover, with the help of context-aware (sensor) technology, learning systems can detect students’ learning behaviors in the real world, and hence, students and educators alike can conduct more active and adaptive learning activities (Hwang, 2006). This learning scenario has been called context-aware ubiquitous learning (context-aware u-learning), which has gradually become a popular trend in education (Ogata & Yano, 2004; Hwang et al., In Press).
In recent years, several research projects have investigated various ways of applying these new technologies to education, and one of the chief goals of these studies has been to strengthen the technologies’ capacity to provide active and adaptive support to real-world learning and training. By accounting for the functions of context-aware ubiquitous computing and interactivity, this study proposes a Ubiquitous Performance-Support System (UPSS) that can facilitate the seamless use of powerful new technologies in the school setting. This study also offers readers one case study that, based on this UPSS, can help them visualize these novel technologies in practice. The aim behind the case study is to inform the design and the development of context-awareness ubiquitous computing system and its learning materials. The research inquiry centers around three themes: (1) the critical features to the data-driven decision making of teachers, (2) the perceptions of teachers and students to the UPSS, and (3) implementation issues. Two rounds of formative evaluation are used to address the three inquiry themes. By presenting the process and the results of the research approach, this study can be a useful reference for those who are interested in conducting studies applying context-aware ubiquitous computing to educational contexts. Finally, this study presents suggestions and implications for future research and system development.

**Backgrounds and Motivations**

In a context-aware u-learning environment, students and educators can conduct real-world learning activities with adaptive supports from the learning system (Rodríguez & Favela, 2003; Ranganathan & Campbell, 2003; Kwon et al., 2005). Several researchers have demonstrated the benefits of such a learning environment in helping the learners to increase their problem-solving abilities in the real world (Jones & Jo, 2004; Hwang, 2006). For example, Ogata and Yano (2004) presented JAPELAS and TANGO, which educators have used to help students learn Japanese under real-world situations. The systems can provide learners with appropriate expressions according to different contexts (e.g., occasions or locations) via mobile devices (e.g., PDA, or Personal Digital Assistant). Rogers et al. (2005) conducted an experiment consisting of indoor and outdoor learning activities. According to the generalized findings, learners can use their observations to gather observation-drawn data, including voice data and image data, and learners can use wireless networks to gather related information from learning activities. Recently, Joiner et al. (2006) presented their studies on education-based use of context-aware devices and noted that the devices quickly offer students vocal statements about real-world conditions. In the meanwhile, Yang (2006) proposed a learning environment that stores resources according to a peer-to-peer (P2P) model and that functions to promote learning-resource sharing.

A performance-support system (PSS) functions to integrate resources into the execution of complex tasks. When these resources include electronic technology, the PSS becomes an electronic performance-support system (EPSS). Both Gery (1991) and Brown (1996) defined an ‘EPSS’ as technology that helps provide users with on-demand access to integrated information, guidance, advice, training, and tools that, in turn, promote high-level job performance with a minimum of support from other people. Therefore, the goal of an EPSS is to provide users with whatever is necessary to ensure performance and learning whenever a user engages in those activities.

The new computational paradigm, which includes ubiquitous computing and the context-aware feature, may create new possibilities for interactivity between humans and computers, and may provide users with supports that are more adaptive than are traditional EPSS supports. The UPSS combines digital and physical resources in novel ways; therefore, rather than see UPSS as merely a vehicle for delivering information, we can see it as a re-conceptualization of the whole work environment—an environment that is grounded in the fluid nature of support in the physical work environment and not in the static nature of formalized knowledge on isolated desktop computers. This concept matches the recent definition that Cagiltay (2006) assigns to EPSS in computer-based systems: “[EPSS] provides support at the moment it is needed (right time), and presents relevant (right type) and context-focused (right amount) information that a task performer needs, in a real work environment (right place)” (p. 94).

Also, the essence of the UPSS is similar to the essence of “mindtool” proposed by Jonassen (2000). Jonassen argues that, in a computer-assisted learning environment, all computer-based tools should be mindtools that function as intellectual partners of learners: in other words, mindtools should facilitate critical thinking and higher-order learning by adapting to the learners (p. 9). And by describing computer-based tools as intellectual partners, he asserts that a given responsibility should “fall to the partner who is better able to perform it” (p. 9). In intellectual partnerships, learners should be responsible for recognizing and judging patterns of information and then for organizing the
information (tasks that humans perform better than computers), whereas computers should perform calculations and store and retrieve information (tasks that computers perform much better than humans).

The school environment that focuses on data-driven decision making has a distinct approach to the capabilities of humans and the capabilities of computers. In this regard, we can envision the possibility that school districts use technology-based solutions to make the handling of mass data more effective, whereas administrators and teachers are the ones to make final decisions based on the data collected from various sources.

A few studies have applied innovative mobile technologies or ubiquitous computing to student learning in fieldtrip settings (e.g., Chen, Kao, & Sheu, 2003), in museum settings (Hsi, 2003; Hall & Bannon, 2006), and in teacher-training settings (Seppälä & Alamäki, 2003). However, there is a dearth of research concerning the novel development and the novel deployment of context-aware ubiquitous computing in relation to data-driven decision-making processes for in-service teachers.

A Conceptual Framework of Data-driven Decision Making and Learning Assistant

Recognizing the importance of using a UPSS both for data-driven decision making and as a mindtool, we used the conceptual framework developed by Mandinach et al. (2006) to guide our design of this study’s UPSS. This framework approaches data-driven decision making as a continuum that stretches from data to information and to knowledge and that addresses cognitive complexity in data where decision-making begins with raw data, transforms those raw data into information, and then ultimately transforms that information into actionable knowledge. The process implies that a successful UPSS needs to become a mindtool for administrators and educators if the UPSS is to perform a series of complex computational tasks before making the best instructional decision. This data-information-knowledge continuum rests on three types of skills: (1) collection and organization (Data Level), (2) analysis and summarization (Information Level), and (3) synthesis and prioritization (Knowledge Level) (Mandinach et al., 2006). Depending on the roles of decision makers and the scopes of school structure (i.e., classroom level, building level, and district level), the data used and the skills needed vary. Further, decision-makers will not always engage these skills in a linear, step-by-step manner. Instead, there will be iterations through the steps, depending on the context, the decision, the outcomes, and the interpretations of the outcomes. Figure 1 displays the procedures and the data types in this data-driven decision-making process.

Figure 1. The conceptual framework for data-driven decision making (Mandinach et al., 2006)
UPSS in Practice: The Butterfly-ecology Course

The research team employed a systematic-design model (Dick, Carey, & Carey, 2005) that focuses on development-implementation-evaluation processes regarding both UPSS-environment development and UPSS-centered supporting materials. While developing the trial version of the UPSS, we conducted two experiments regarding the science course “Butterfly Ecology” at an elementary school in Taiwan: the first experiment took place from July to August 2007 and the second experiment, from January to February 2008. This particular school has been known for its butterfly-ecology training program and has been certified to regularly offer training programs for garden guides (Chu et al., In Press). Owing to the explorative nature of the study and the limited budget, we used two rounds of cost-effective evaluative strategies (e.g., expert review and two field trials) to provide an immediate tactical analysis of the function feasibility and implementation issues in the real-world setting. By presenting the process and the results of our evaluative strategies, we hope to provide a guide to others who are interested in applying context-aware ubiquitous computing to educational contexts.

To be more specifically, three themes were identified for analysis in this study.

1. What are important features of UPSS are critical to the data-driven decision making of teachers?
2. What are teachers’ and students’ perceptions of UPSS, with regard to motivation and learning?
3. What implementation factors are identified when considering the feasibility of integrating the UPSS into the existing curriculum?

The Context and the Participants

This study conducted two rounds of evaluations: round one included one instructional designer and one subject-matter expert (SME) and five elementary school students, and round two included nine science teachers and thirty elementary school students. These students took a special training program about butterfly-ecology gardens. The most challenging issue during the training was the mastery of observation and classification skills. To be sure, this process requires more than simply memorizing the appearance of plants or the facts from the textbook. In a traditional class in this field, a trainee teacher needs to deal with ten or more students and to develop the students’ skills in observation and classification. With such a one-to-many approach, it is difficult for the trainee teachers to provide individual instruction and, in particular, to make informed decisions based on each individual student’s feedback. Figure 2 shows one corner of the garden, which consists of 25 ecological areas with different plants and butterflies. The research team attached a radio frequency identification (RFID) tag containing key information on each target plant. Moreover, the knowledge database ran on a mobile device and enabled the students to tour a particular plant and the species of butterflies. While the students moved throughout the garden, via wireless communications, the UPSS displayed their learning portfolio consisting of students’ locations and learning progress in relation to the garden, detected the information from the tags, and sent prompts to the students’ mobile devices. Once the students entered their responses on the mobile device, the UPSS stores and aggregates the information from the embedded sensors connected to the networks (This process is considered as Data Level in our design framework).

In order to facilitate the trainee teachers’ data-driven decision-making process, the research team also incorporated expert advice into the UPSS for problem structuring, decision support, analysis, and diagnosis. This advice existed in many forms that trainee teachers could invoke according to their needs. The advice involved an expert system that asked a trainee teacher questions and that then suggested the most appropriate procedure or step in response to the questions (Information Level). Or the advice involved an interactive expert system that used case-based reasoning or coaching to guide the teachers through decision-making processes (Knowledge Level). Table 1 tabulates the design concepts of the UPSS functions in the data-information-knowledge continuum.
Table 1. The UPSS functions in the data-information-knowledge continuum

<table>
<thead>
<tr>
<th>Level</th>
<th>The UPSS Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>The UPSS actively collects and organizes students’ data into learning portfolios (i.e., students’ touring locations, their responses to the prompts provided by the UPSS) from the embedded sensors at the garden. The trainee teacher can make use of data in the fastest possible time and with a minimum of manually entered data.</td>
</tr>
<tr>
<td>Information</td>
<td>The UPSS summarizes and reports data to the trainee teacher. It focuses on approaches to analyzing and selecting needed information, on available resources, and on relating new knowledge to existing knowledge and experience.</td>
</tr>
<tr>
<td>Knowledge</td>
<td>The UPSS alerts the trainee teacher to available tools and resources that may be helpful under given circumstances. The UPSS also promotes the trainee teacher’s contemplation skills by asking questions and by prompting the trainee teacher to identify a similar learning event on the basis of his or her own experiences.</td>
</tr>
</tbody>
</table>

The Trainee Teacher Interface

There are two types of UPSS interfaces: one is designed for the students and the other is for the trainee teachers. Figure 3 presents the UPSS trainee-teacher interface, which provides the trainee teachers with a tabulated interface that helps them not only maintain the course content and the knowledge base but also browse the profiles and the learning portfolios of the students. Each mobile device is equipped with an RFID reader, which can detect the signals sent from the RFID tags attached to each target object (e.g., the plant). The signals contain the unique identification codes of individual targets. As the students move around the garden, the UPSS can automatically detect, store, and aggregate the information as it moves from the embedded sensors and through the wireless networks. As shown in the right section of Figure 3, records in the portfolio database consist of several data fields on students’ learning behaviors: locations, on-line learning materials browsed, on-line requests for help, notes taken, responses to questions, content of on-line discussions, duration of each observation, and each student’s location-arrival time.

The Student Interface

Figure 4 displays the student interface on the mobile device. In the left screenshot, the UPSS guided the student to a corner of the butterfly garden and asked the student to find a plant called “Birthwort.” After the student walked close to the plant, the UPSS displayed detailed information about the plant, as shown in the middle screenshot. In the right screenshot, the UPSS guided the student to the butterfly larva on that plant.
Figure 3. Teacher interface of the UPSS

- Learning portfolio
- Courseware management
- Knowledge base management
- Account management

Digital content browsed
Using “Help” function (Yes/No)
Using “Notebook” function (Yes/No)
Test records
Time spent in the location
Time for browsing the digital content
Conversation target (time)

The system guided the student to a butterfly ecology area.
The system introduced the features of the plant when the student arrived at the area.
The system guided the student to observe the candidate butterflies in that area.

Figure 4. Screenshots of the UPSS observation functions for butterfly ecology
Data Collection and Analyses

The system-development process was recursive, going through several phases of development, testing, and revision according to procedures in the two rounds of formative evaluations (see Dick et al., 2005): (1) expert review of both content and interface design by one expert in instructional design and by one SME expert, and (2) two field trials by potential users, including teachers and students. To address the three proposed inquiry themes, the research team collected both qualitative data and quantitative data from teachers and students over the two six-week implementation periods. The primary purpose of round one was to evaluate the accuracy and the currency of the training content and to evaluate the function feasibility of the UPSS (research inquiry one). The primary purpose of round two was to detect the implementation issues surrounding teachers’ application of UPSS to school settings (research inquiry two and three). Therefore, the experts involved in round one were the ones who had designed the training content and had worked closely with the system programmers regarding the functions; the teachers and the students in round two were the potential users who might use the UPSS in their courses. At the end of the two experiments, we conducted surveys and interviews with the teachers and the students about their perceptions of the UPSS. The results informed our understanding of how the UPSS worked; and in turn, we sought to refine our system design.

In addition to the datasets that we collected from surveys and interviews, the UPSS simultaneously collected and organized data regarding the “physical” learning behaviors of individual students. These types of behavior concerned time, location, and student-system interaction (e.g., how, and after their observations, individual students responded to the system’s prompts and instructions). The UPSS also enhanced students’ reflections by comparing the students’ responses with pre-defined learning behaviors or with expected answers (provided by the trainee teachers). A knowledge base embedded in the system rigorously calculated and prioritized critical misconceptions of individual students and helped enhance individual students’ reflections.

For example, assume that UPSS is trying to guide the student to find “Birthwort” in the butterfly garden; nevertheless, the student walks to a wrong location and show a different plant to the system. UPSS will invoke the knowledge base to find the possible misconception, and then lead the student to make reflection by showing some hints and conducting the student to compare the differences between the two plants. Figure 5 shows the interface for supporting the student to make reflection while encountering difficulties in recognizing the target plants.

Figure 5. Screenshots of the UPSS for providing reflection supports
**Round One: Expert Review**

In round one, the draft version of the UPSS was reviewed by one instructional designer and one SME who operated outside the project and who had special expertise in the content area of the training. They were to comment on the accuracy and the relevance of the system and the training program. They also critiqued the function feasibility whose design rested on UPSS concepts and on the framework of data-driven decision making. This specific SME was also familiar with the target population and perhaps could provide insight into the appropriateness of the learning materials for the eventual performance context. The two experts actively observed five elementary students’ use of the UPSS and engaged these students to contribute their ideas in the design process. Thereafter, the research team conducted in-depth interviews with these experts and the students. After collecting and summarizing the experts’ data and the students’ input, the research team made minor changes to the draft version of the UPSS and continued the second round of evaluation.

The review dimensions by the instructional designer and the SME included “Clarity” (“Are the training materials clear to the individual target students?”), “Effect” (“What effect does the instruction have on individual students’ attitudes and on individual students’ achievement of the objectives?”) and “Feasibility” (“How feasible are the training materials in relation to the availability of resources such as time and context?”). Accordingly, the research team interviewed the trainee teacher and the SME with the following questions:

1. How long (in years) have you been training students in the butterfly-ecology program?
2. How many hours per week do you usually teach with the butterfly-ecology training materials? Do you use any instructional tools to facilitate your training process?
3. What is the average timeline that characterizes your preparation of a skilled guide for the butterfly-ecology garden?
4. Are there situations in your use of the UPSS that are unique or different from situations that arise in traditional training programs? If so, what are the situations?
5. What are a few important features that a UPSS should have? What tools or features might be added to the UPSS to make it better for your training? For instance, what types of activities or tools would you suggest to include in the UPSS in order to promote individualized learning?
6. What costs and benefits, if any, do you associate with using the UPSS for the training?
7. What is your perception of this alternative training mode’s motivation of your students or the mode’s enhancement of students’ learning?

Furthermore, the students were interviewed with the following questions:

1. Have you attended a similar butterfly-ecology class before?
2. Was the butterfly-ecology program interesting?
3. How do you feel about using the UPSS for this training? Are there any differences between use of the UPSS and traditional training programs? If so, what are they?
4. What would you tell other students about the UPSS project?

**Round Two: The Field Trial**

The field trial used a learning context that closely resembled the learning context intended for the actual applications of both the UPSS system and related learning materials. One purpose of the round-two evaluation was to determine whether or not the post-review changes that the research team had made in both the system and the learning materials were effective. Another purpose was to determine whether or not the instruction was applicable to its intended context. A total of nine teachers and thirty elementary-school students participated in the field-trial rounds. In round two, the research team (1) interviewed nine teachers regarding both the integration of the UPSS into their science course and their perceptions of the benefits or the limitations therein, and (2) asked thirty students to fill out a survey regarding their experiences using the system. Here are the questions that the research team interested:

1. What are the differences between your use of the UPSS and your regular teaching practices?
2. What, if any, teaching benefits do you attribute to your use of the UPSS, particularly regarding teaching-load dimensions, students’ motivation, and learning effectiveness?
3. Do you anticipate any costs that may arise from school-based use of the UPSS?
4. Will you try the UPSS for your science class in the future? If so, what pedagogy will you use? If not, why not?
5. Will you recommend the UPSS either to teachers or for subjects outside the natural sciences?
Results

After collecting the data from the two rounds of evaluations, the research team examined usability, student motivation, student learning, and classroom implementation issues, as shown in the following subsections.

Issues of Motivation, Management, and Effectiveness

Expert Review

Two experts were interviewed following the round one evaluation. Expert one (a science teacher, coded E1) had five years of experience in conducting the butterfly course, and Expert two (a science teacher, coded E2) had eight years of experience therein. When asked to comment on the differences between the ubiquitous learning environment and the traditional training process, the teachers voiced similar points of view: the teachers stated that both themselves and students were greatly benefited the mobile-learning environment in terms of motivation, manageability, and effectiveness.

For example, in relation to the “motivation” aspect, E1 stated, “It is obvious that the use of mobile devices can motivate the students. The multimedia materials are very attractive.” She particularly favored the interactive-learning program: “In contrast to the traditional instructions given by the teacher, the students seemed to be happy while interacting with their mobile devices, which enhanced their motivation levels in the training.” Further, E2 mentioned the “continuous concentration” of the students that manifested itself while the mobile device guided them during the learning process. He stated, “It is amazing that the students concentrated on the learning activities for a longer time than I ever expected.”

As to the “manageability” aspect, E1 and E2 agreed about the benefits of the UPSS. E1 stated, “With this innovative UPSS system, we can browse the learning process of individual students. This would allow me to keep track of the real-world learning status of each student.” E2 stated, “By recording every action and the paths of the students, the computer helps review their learning activities, and this significantly helps adjust the learning on the basis of the UPSS-collected data.”

When asked about the functionality of the UPSS-provided guidance, E1 positively regarded the “effectiveness” aspect. E2 responded in kind: “The mobile device can correctly guide [students to] or hint at the learning paths and can, therefore, provide adequate learning content to students. It works like a tutor for each student, and I think most of the students would benefit from this UPSS system.” Thus, it is reasonable to conclude that both the trainee teachers and the students considered the UPSS useful.

Student Feedback

There were two major sources of students’ feedback: one was from the round one and the other was from the round two. In round one, five students, coded as student one (S1), S2, S3, S4, and S5, were selected for one-to-one semi-structured interviews. At the time, all these students were fifth graders and were eleven years old. They had taken a basic course concerning butterfly ecology. The chief purpose of the interview was to elicit UPSS-related responses and suggestions from the students after they had experienced the trial version of UPSS. All the interviews were recorded on a digital recorder, and all the data presented in this paper were then analyzed and translated by the authors.

When asked to discuss the differences between the UPSS and the traditional training process, the five students variously characterized the UPSS as an “interesting,” “interactive,” or “effective” learning assistant. For example, S1 found the UPSS “interesting and interactive” in comparison with the traditional training; S1 stated, “Unlike traditional training processes by teachers, the UPSS is more interactive because it shows every detail, which motivates my learning.” He favored the interactive learning program: “I can go over the learning path repeatedly with the mobile device, rather than ask repetitive questions to the teachers. This flexibility has relaxed me in the learning process.” S2 raised similar points: “The mobile device guides my observation of butterfly ecology step by step, and this makes me feel like I have a patient tutor who, on the side, reduces the learning pressure.” The above
responses indicate that the UPSS can provide friendly interface and a friendly learning environment. It is worth noting that, according to both S4 and S5, their use of the mobile device enabled them to access data when needed, a function that heightened their interest in the course.

As to the “effectiveness” perspective, all the students viewed the UPSS positively. For instance, when asked about how the learning-guidance function worked, four students (S1, S2, S3, and S4) noted the immediacy and the convenience of the system. S3 commented, “The mobile device can be an accurate guide or can suggest a learning path, and this makes me feel like I have a teacher beside me. I would prefer learning with this system to learning with teachers.” Thus, it is reasonable to conclude, also, that the UPSS had a positive effect on the motivation of these students.

In round two, thirty students were asked to fill out the after-project survey. The three dimensions surveyed in this stage are students’ prior knowledge and skills in technology and butterfly ecology, experiences of learning, and students’ overall impression of the unit. The items of these dimensions were presented on a 5-point Likert scale, ranging from “strongly disagree” (1 point) to “strongly agree” (5 points). The results demonstrated students’ positive perceptions of the learning unit. Table 2 shows the results of this field trial of this learning unit.

<table>
<thead>
<tr>
<th>Dimensions and evaluation items</th>
<th>N</th>
<th>A</th>
<th>NE</th>
<th>D</th>
<th>SD</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>I observed the plants and butterflies in the butterfly-ecology garden before the teacher introduced this learning unit.</td>
<td>17</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4.47</td>
</tr>
<tr>
<td>I used the personal digital assistant (PDA) before this learning unit.</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>13</td>
<td>4</td>
<td>2.63</td>
</tr>
<tr>
<td>I used a digital camera for butterflies before this learning unit.</td>
<td>11</td>
<td>14</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>4.20</td>
</tr>
<tr>
<td>I searched the Internet for information on butterflies.</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>3.07</td>
</tr>
<tr>
<td>My teacher used multimedia (e.g., images or videos) in class before this learning unit.</td>
<td>15</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4.40</td>
</tr>
</tbody>
</table>

**Part II. Learning Experiences**

| I enjoy observing butterflies because of their attractive appearances. | 7 | 12 | 7 | 4 | 0 | 3.73 |
| I enjoy observing butterflies because I enjoy a natural life. | 6 | 16 | 5 | 3 | 0 | 3.83 |
| I enjoy observing butterflies because I can gain natural knowledge. | 5 | 12 | 10 | 3 | 0 | 3.63 |
| I was less motivated by my teachers’ tour and lectures than by the illustrations and the guidance of the UPSS during this learning unit. | 19 | 8 | 3 | 0 | 0 | 4.53 |
| I tried to search information via the UPSS in this learning unit. | 4 | 10 | 8 | 8 | 0 | 3.33 |
| I think I have gained knowledge with the guidance of the UPSS in this learning unit. | 11 | 8 | 9 | 2 | 0 | 3.93 |
| I felt that using the UPSS made it convenient to observe butterflies in this learning unit. | 15 | 15 | 0 | 0 | 0 | 4.50 |
| I felt that my learning with the UPSS was easy in this learning unit. | 16 | 14 | 0 | 0 | 0 | 4.53 |
| I was more motivated in my discussions with classmates when I was using the UPSS than when I wasn’t using the UPSS in this learning unit. | 8 | 11 | 6 | 5 | 0 | 3.73 |
### Part III. Overall Impressions

<table>
<thead>
<tr>
<th>Dimensions and evaluation items</th>
<th>SA N (%)</th>
<th>A N (%)</th>
<th>NE N (%)</th>
<th>D N (%)</th>
<th>SD N (%)</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have learned about the function of PDAs in this learning unit.</td>
<td>12 (40.0)</td>
<td>15 (50.0)</td>
<td>3 (10.0)</td>
<td>0</td>
<td>0</td>
<td>4.30</td>
</tr>
<tr>
<td>I have learned by observing through the UPSS in this learning unit.</td>
<td>10 (33.3)</td>
<td>18 (60.0)</td>
<td>2 (6.7)</td>
<td>0</td>
<td>0</td>
<td>4.27</td>
</tr>
<tr>
<td>I have learned about butterflies in this learning unit.</td>
<td>7 (23.3)</td>
<td>18 (60.0)</td>
<td>5 (16.7)</td>
<td>0</td>
<td>0</td>
<td>4.07</td>
</tr>
<tr>
<td>I feel that my interest in butterflies has grown since I completed this learning unit.</td>
<td>11 (36.7)</td>
<td>14 (46.7)</td>
<td>5 (16.7)</td>
<td>0</td>
<td>0</td>
<td>4.20</td>
</tr>
<tr>
<td>I will try to search for more information on butterflies after the class.</td>
<td>4 (13.3)</td>
<td>14 (46.7)</td>
<td>7 (23.3)</td>
<td>5</td>
<td>0</td>
<td>3.57</td>
</tr>
<tr>
<td>I felt that learning via the UPSS was much more interesting than learning via my teacher’s tour and lectures.</td>
<td>16 (53.3)</td>
<td>14 (46.7)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.53</td>
</tr>
<tr>
<td>I am satisfied with learning with the UPSS.</td>
<td>17 (56.7)</td>
<td>13 (43.3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.57</td>
</tr>
<tr>
<td>I felt that it is easier to learn with the UPSS (one student per PDA) than with my teacher’s tour and lectures.</td>
<td>12 (40.0)</td>
<td>18 (60.0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.40</td>
</tr>
<tr>
<td>I think I can learn more with the UPSS than with my teacher’s tour and lectures.</td>
<td>15 (50.0)</td>
<td>11 (36.7)</td>
<td>4 (13.3)</td>
<td>0</td>
<td>0</td>
<td>4.37</td>
</tr>
<tr>
<td>I think I am more comfortable learning with the UPSS than with my teacher’s tour and lectures.</td>
<td>9 (30.0)</td>
<td>16 (53.3)</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>4.13</td>
</tr>
<tr>
<td>I think I am more motivated to learn with the UPSS than with my teacher’s tour and lectures.</td>
<td>12 (40.0)</td>
<td>16 (53.3)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4.33</td>
</tr>
<tr>
<td>I will recommend this learning unit to other classmates.</td>
<td>9 (30.0)</td>
<td>12 (40.0)</td>
<td>6 (20.0)</td>
<td>3</td>
<td>0</td>
<td>3.90</td>
</tr>
</tbody>
</table>

SA = Strongly agree, A = Agree, NE = Neutral, D = Disagree, SD = Strongly disagree

The survey also asked how students felt about using the UPSS in the learning unit. Twenty-nine out of thirty students expressed that they had enjoyed using the UPSS and the personal digital assistant (PDA) for butterfly observation. For example, most students found use of the UPSS to be motivating because the system allowed for more self-paced learning and for more comparisons with the teacher’s tour and lectures. Two students liked the convenience of immediate information retrieval, while one student mentioned that the system would provide personal guidance based on his inquiries. Only one student mentioned a technical problem (e.g., disconnection, and double-clicking for correct answers) that had hampered his learning.

### Implementation Issues

In round two, the research team conducted one-to-one semi-structured interviews with nine teachers, who were coded as teacher 1 (T1) to teacher 9 (T9) and who had used the UPSS in their classes. All these teachers were trained in the sciences subject, and the teachers’ years of teaching experience varied from five to sixteen years.

From the teacher interviews, the research team found that teachers’ attitudes toward and teachers’ pedagogical approaches toward UPSS use are two key factors in the successful implementation of teaching with innovative technology. With regard to the teachers’ attitudes, most teachers had positive attitudes toward UPSS use and were willing to adopt it in their future teaching practices. Round-two findings were similar to round-one findings: these teachers agreed on the advantages of higher student motivation, as well as the individualized and adaptive supports resulting from teachers’ adopting the UPSS. For example, they mentioned that learning with UPSS created many possibilities that were not easily achieved in the traditional learning environments. Some of these possibilities were (1) students’ better access to mobile computers and to online resources and (2) reduced constraints relative to time or
location (T3, T6, T8, and T9). Meanwhile, teachers believed that learning activities with the UPSS could help engage students in learner-centered activities seamlessly from one physical location to the next: part of this seamlessness stemmed from the UPSS’s powerful ability to transfer data from the ecology garden to the computer lab and vice versa. Another strength of the UPSS lies in its data-collection features, its learning-portfolio features, and step-by-step expert advice which are automatically stored and aggregated through sensors and networks (T2, T3, T4, T5, T6, and T9). Indeed, T3 mentioned that “these contextual data are particularly useful for individualized or supplementary instruction after the learning activity, and these data can be another eye for the teacher.” However, two teachers (T8 and T9) contended that they might not have time to create a UPSS-related learning activity because of tight schedules, particularly regarding the time constraints imposed by term exams. T9 doubted that some teachers at school have enough background knowledge of, or experience with, mobile computers to teach with such advanced technology; for example, many teachers had never used a personal digital assistant. Such feedback helped the research team confirm that, before successful implementation takes place, teachers must undergo sufficient technology training and must have access to handy materials—all grounded on sound learning theories and the features of context-aware ubiquitous computing.

These teachers believed that the UPSS can be implemented in many disciplines—such as the natural and pure sciences, the social sciences (e.g., tours of historical sites), and museum education—as long as the implementation takes place under student-centered pedagogical approaches, such as situated learning, inquiry- or problem-based learning, and collaborative learning. For instance, T4 described a learning scenario anchored around a theme of cultural or historical relics. He expected his role to be that of either a guide or a facilitator, and he expected to promote opportunities in which his students could initiate and control their own learning. Furthermore, he suggested that teachers who adopt not standard tests but innovative technology such as a UPSS should consider instructional goals and should accompany the technology with alternative assessment methods, such as portfolio evaluations, tokens, or group competition. In this way, he stated, the teachers could evaluate students’ learning processes and learning outcomes. He added, “The UPSS collects rich data that the teachers should not overlook.”

**Conclusions**

The advance of computer and network technologies has encouraged researchers to conduct various learning activities with mobile devices and wireless communication. This paper proposes a ubiquitous performance-support system (UPSS) that can facilitate the seamless use of powerful new technologies in school settings. On the basis of innovative approaches, we developed a ubiquitous-learning system for a butterfly-ecology course by systematic-design approach to refine the system and its supporting materials. By engaging in two rounds of design, implementation, analysis, and re-design, we have been able to refine both our curriculum and the UPSS environment prior to conducting formal randomized experimental trials in the future.

After the two rounds of evaluations, several parts of the design (e.g., step-by-step expert advice for teachers to generate reports based on students’ portfolio data) have remained relatively unchanged from the initial implementation because our analysis indicated that these parts were successful in meeting our objectives. Other parts of the design have changed on the basis of feedback from round one. For example, we increased the number of tacit clues embedded in the UPSS to ensure continuous engagement of high-performance students in materials provided in the curriculum.

The results of formative evaluations reveal that the experts and students had positive perceptions of the UPSS in relation to motivation, interactivity, and effectiveness; in addition, both of the experts agreed that the mobile device’s learning guidance (in the context of the real-world garden) was essentially helpful to the students. These reactions point to the effectiveness of this innovative training approach. Furthermore, as indicated in round two, the UPSS, because of its data-driven decision-making functions, enabled teachers to keep track of the students’ learning status as well as facilitated the teachers’ individualized instruction. Overall, these findings promote further refinement of and experimentation with UPSS as a training modality that can help teachers enhance students’ motivation. These data are promising but not conclusive about the UPSS’s educational value. By examining the fact that students found the UPSS readily usable and the learning experiences motivating, we also found weaknesses in the design, particularly from a graphical and curricular perspective. For example, teachers need more support for implementing this UPSS into their curriculum. As a result, we proposed a rigorous training program of professional development
for teachers. The program functions, in part, to familiarize teachers with a constructivist pedagogical approach that enables students to actively and collaboratively explore the domain knowledge.

In traditional training settings or classrooms, teachers usually have to use their judgment or tests to gauge students’ learning processes and learning outcomes after an experiment. Under such conditions, teachers often overlook much of the valuable contextual information; moreover, teachers must store and analyze this information in order to create individualized supports that, by accurately reflecting students’ actions, enhance them. In the round-two interviews, the teachers confirmed the advantages of using context-aware ubiquitous computing as a mindtool (the data-collection and data-driven tool): such computing helped the teachers focus on decision-making practices. This finding also has an important implication for researchers: context-aware ubiquitous computing can serve as an objective yet unobtrusive data-collection tool depicting students’ learning processes, whereas tests or surveys capture a much smaller wedge of the spectrum of students’ learning, especially when students are engaged in an innovative-technology learning environment.

In conclusion, we assert that the innovative system proposed in this study can benefit teachers, students, as well as researchers. Future studies may examine the relationships between learner perceptions and learner usage of the UPSS and learners’ overall satisfaction, in a larger scale, with longer experiment periods. Also, it may be interesting to explore the teachers’ perceptions of using UPSS and their instructional strategies as implementing the UPSS into their existing curriculum.

Acknowledgements

This study is supported in part by the National Science Council of Taiwan, the Republic of China, under contract numbers of NSC95-2520-S-009-009, NSC NSC 97-2631-S-024-002, and the Taiwan Ministry of Education (project title: ATU Plan of the National Chiao Tung University and Ministry of Education, Taiwan, R.O.C.).

References


Content-Free Computer Supports for Self-Explaining: Modifiable Typing Interface and Prompting

Chih-Yueh Chou and Hung-Ta Liang
Department of Computer Science and Engineering, Yuan Ze University, ChungLi, Taiwan // Tel: +886-03-463-8800 ext 2362 // Fax: +886-03-463-8850 // cychou@saturn.yzu.edu.tw

ABSTRACT
Self-explaining, which asks students to generate explanations while reading a text, is a self-constructive activity and is helpful for students’ learning. Studies have revealed that prompts by a human tutor promote students’ self-explanations. However, most studies on self-explaining focus on spoken self-explanations. This study investigates the effects of two kinds of content-free computer supports for self-explaining: one is full-text reading and modifiable typing interface and the other is prompting. The results showed that self-explaining in a full-text reading and modifiable typing interface, particularly when prompted, made students perform better in applying target procedural knowledge to similar problems than reading did. The results also showed that typing self-explaining can make similar high-quality self-explanation proportion to spoken self-explaining. In addition, this study investigates the effects of content-free prompts by comparing with no prompts and with content-related prompts. This study uses the computer as a learning companion to provide prompts without understanding the content of students’ self-explanations. The results showed that both content-free and content-related prompts made students generate more self-explanation modifications and perform better in applying target procedural knowledge to similar problems.

Keywords
Computer assisted learning, Self-explaining, Learning companion, Interactive learning environment

Introduction
Self-explaining, which engages the student in explaining to him/herself while studying a text or while solving problems, has aroused much attention since 1989. In a research of having students spontaneously explain the text, students were classified as good students and poor students according to their posttest performance (Chi et al. 1989). The research showed that the good students spontaneously generated more self-explanations than poor students did. In a research of comparing a group of students under spontaneously self-explaining condition and a group of students under self-explaining with prompts from a human tutor, prompted students performed better than unprompted students did (Chi et al. 1994). However, individual difference exits when received prompts; that is, some prompted students generated more self-explanations and performed better in posttest than some other prompted students did. In addition, some unprompted students spontaneously generated many self-explanations. To account for the self-explaining phenomenon, Chi (2000) pointed out that self-explaining is a self-constructive activity that involves generating inferences to fill in information omissions in the text, as well as monitoring and repairing faulty knowledge. The inferences can be generated by integrating information presented across different sentences, by integrating information with prior knowledge, or by using the meanings of words to imply what may also be true. Knowledge repairing is a process of revising one’s own imperfect mental model. The students' self-explanations can be classified as re-reading, paraphrase, inference, self-monitoring, etc. (Chi, 2000; McNamara, 2004). In addition, re-reading and paraphrase could be noted as low-quality self-explanations (LSE) and inference and self-monitoring could be noted as high-quality self-explanations (HSE) (Roy & Chi, 2005). Studies also revealed that HSE is positive related to learning gains (Chi & Bassok, 1989; Pirolli & Recker, 1994). In fact, in many studies, only HSE is regarded as self-explanation.

Most research on self-explaining engaged students in speaking out explanations while reading a text. However, some research applied computers to support self-explaining. For instance, one study provided a menu-based interface for students to generate domain-based self-explanations through selection from a set of domain-based rules or plans (Conati & VanLehn, 2000a). The results showed that some self-explaining students performed better than other students without self-explaining, but some self-explaining students performed worse. The possible reason might be that self-explaining is a self-constructive activity and choosing from a menu of explanations is not so constructive (Hausmann & Chi, 2002). In contrast with the menu-based interface, another research project allowed students to generate explanations by typing on the keyboard when solving problems (Aleven & Koedinger, 2002). The results
showed that self-explaining through a typing interface also makes students perform better at problem solving. Another example of applying a computer in self-explaining was to investigate whether a computer interface can support self-explaining (Hausmann & Chi, 2002). In the research, the students read a text sentence by sentence and typed their explanations on each sentence. The findings showed that the spoken explanations tended to be fragmented and incoherent while typed explanations could be noted for their completeness. The results also revealed that students generated fewer explanations in a typed form than that in a spoken form. Several possible reasons were proposed by researchers to account for the results: typing requires more cognitive capacity and resource than speaking, and typing provides a record so that students avoid errors. In addition, the research found that content-free prompts also benefited students’ learning even in a typed form self-explaining environment. The content-free prompts are prompts without any domain-related or content-related information that can therefore be easily adopted in any domain or for any content, such as “Could you elaborate on what you just said?”

![Diagram](image)

**Figure 1.** Interface One: reading

However, many issues surrounding the use of a computer to support self-explaining remain to be investigated. For example, Hausmann & Chi (2002) investigated the effects of using a typing interface for self-explaining the text sentence by sentence, but students could not re-read the previous sentences or change their previous self-explanations when reading later sentences in the text. However, students might want to re-read some sentences or have different understandings on some sentences when they read later sentences. Typing records might also provide an opportunity for students to reflect on or revise their understanding of the text. Therefore, the first research issue we want to investigate is 1) the effects of self-explaining in a full-text reading and modifiable typing interface. The interface allows students to freely read the whole text according to their chosen sequences; that is, students can read, skip, or re-read some sentences. The interface also allows students to type self-explanations and modify their
previous self-explanations. The study aims to investigate learning performance and self-explaining behaviors in such an environment. We are also interested in the students’ reading and self-explaining sequences, and whether and how students modify their self-explanations when they are allowed to modify them.

The second research issue we want to investigate is the effects of different kinds of prompts in self-explaining. The content-free prompts are showed to be helpful in promoting students to generate more self-explanations and to learn better (Hausmann & Chi, 2002). A research of comparing generic content-free prompts and specific content-related prompts revealed that the less able students learned better with specific prompts and the more able students learned better with generic prompts (Aleven et al. 2006). However, the detailed effects of different prompts in self-explanation remain unclear. In a full-text reading and typing self-explaining environment, this study aims to examine the effects of different kinds of prompts in self-explaining. Instead of providing adaptive prompts by human tutors, this study uses a learning companion as a prompter. The learning companion is a computer-simulated character that plays a non-authoritative role and may provide incorrect information (Chan & Baskin, 1990; Chou et al. 2003); that is, the prompts are designed in advance and are provided for specific sentences of the text without understanding the content of students’ self-explanations.

Figure 2. Interface Two: reading and self-explaining
System

To investigate the two research issues, a system was implemented with three kinds of interfaces: reading (Interface One), reading and self-explaining (Interface Two), and reading and self-explaining with prompts (Interface Three). Interface One is a full-text masking reading interface. The interface divides the text into several text fields, with one field containing one sentence or one diagram (Figure 1). The student reads the text by moving the mouse from one field to another field. When the student moves the mouse into one of the text fields, the sentence is revealed. When the student moves the mouse out of the text field, the sentence will disappear. The student can skip some sentences or re-read any sentence. The masking interface can help students focus their attention and allow the system to trace the student’s attention (Conati & VanLehn, 1999; 2000b). In this study, the masking also allows the system to trace the sequence of how the students read and self-explain.

Interface Two contains a similar masking reading interface to Interface One, and, additionally, there is a self-explanation field beside each text field for students to type their self-explanations (Figure 2). However, in Interface Two, when a student types in a self-explanation field, the sentence of the corresponding text field is revealed. Therefore, in Interface Two students can read two fields at the same time if they put the cursor in a self-explanation field and move the mouse to another text field. Interface Two allows the students to read the text and generate their self-explanations by typing. Furthermore, the students can skip generating self-explanations on a sentence or modify self-explanations on previous sentences. Thus, the modifications of the self-explanations in each field can be recorded and analyzed. The masking mechanism was only applied to text fields, and was not applied to self-explanation fields so that students could easily know which field was being explained and read or modify their previous self-explanations.

Figure 3. Interface Three: reading and self-explaining with prompts
The reading and self-explaining interface of Interface Three is similar to that of Interface Two, but Interface Three can display prompts on some fields (Figure 3). In contrast with the system of providing prompts after students read and self-explain a sentence that was used in the study of Hausmann and Chi (2002), Interface Three allows students freely to read and self-explain the whole text for a period of time, and then provides prompts on some sentences. After receiving prompts, the students can continue to read and self-explain the text. In this study, the prompts are designed in advance to promote students’ self-explanations and are provided to students without understanding the content of students’ self-explanations.

Experiment

An experiment was conducted to investigate the effects of self-explaining in a full-text reading and typing self-explaining interface and the effects of different kinds of prompts for self-explaining. The participants were 75 college undergraduate students enrolled in a Computer Programming course and majoring in computer science. The participants had learnt binary tree in a previous Basic Computer Concepts course.

Materials

The text is to introduce the concept of a red-black tree and to teach the building process of a red-black tree. A red-black tree is a kind of binary tree with some specific limits. The building process of a red-black tree involves several steps: inserting a node according to the definition of a binary tree, checking whether the insertion makes this tree
beyond the limits of a red-black tree, and transforming the tree to fit the limits. The first part of the text introduces conceptual knowledge of a red-black tree; that is, it describes the definition and specific limits of a red-black tree. The second part of the text presents procedural knowledge of a red-black tree; that is, it uses a set of diagrams to present an example about how to build a red-black tree with six nodes containing data from 1 to 6 from one node to six nodes, node-by-node and step-by-step (Figure 4). The second part does not contain other text clarifications about the building process of the example. Thus, the text omits some information, such as checking the status of the limits and the principles of transforming a tree. The design of the text aims to observe whether the students infer omitted information in the text or not. The text was divided into 21 fields and was put into a masking reading interface as shown in Figure 1. The first part of the text contains eight fields and the second part of the text contains 13 fields.

The pretest includes a Binary Tree Definition Test (BTDT) and a Red-black Tree Definition Test (RTDT). BTDT has three questions and each question shows a tree for students to judge whether the tree is a binary tree or not. The student can answer “Yes”, “No”, or “I don’t know”. Each question was scored as 1 if the student answered correctly and as 0 if the student answered incorrectly or answered “I don’t know”; that is, full marks for the BTDT is 3. Similarly, RTDT has three questions for students to judge whether the tree is a red-black tree or not. The pretest is designed to assess the students’ ability in binary tree (prior knowledge) and red-black tree (target knowledge). The posttest includes a BTDT, an RTDT, a Retention Test, and a Transfer Test. The questions in the BTDT and RTDT are the same as those in the pretest for assessing whether students could classify these trees; that is, assessing conceptual knowledge according to Bloom’s taxonomy (Anderson et al. 2000). The Retention Test asks the students to draw the building process of a red-black tree with six nodes as the same as the text and aims to assess whether the students understand and remember the building process (procedural knowledge) of a red-black tree in the text or not. The Transfer Test assesses whether the students are able to apply the knowledge to another similar problem (Bransford et al. 2000). The Transfer Test includes two questions. The first question asks the students to draw the building process of a red-black tree with six nodes of 3, 6, 1, 5, 4, and 2. The content of these nodes is the same as that of the text, but the sequence of these nodes is different and it makes the building process different. The first question is termed as the Near Transfer Test. The second question is to add a node of 16 into a red-black tree with nodes of 10, 20, 30, 15, 13, 14, and 17. The second question involves several transformations and thus is more difficult than the first question. The second question is termed as the Far Transfer Test. In the Retention Test and each question of Transfer Test, the students’ answers were scored as a decimal fraction from 0 to 1 according to their correctness. As an example, there are about 10 steps (some steps can be combined) in the answer to the Retention Test. Thus, the students got 0.1 for each step they performed correctly.

Grouping and procedure

Participants were divided into four groups: Reading (termed as Group R), Self-Explaining with No Prompts (Group SENP), Self-Explaining with content-Free Prompts (Group SEFP), and Self-Explaining with content-Related Prompts (Group SERP). In investigating the issue of the effect of self-explaining in a modifiable typing interface, Group R is a control group and Group SENP is an experimental group. Using the same modifiable typing interface to generate self-explanations, Group SEFP and Group SERP could also be regarded as experimental groups on this issue, although the prompting effect was involved. In exploring the effects of different kinds of prompts in self-explaining, Group SENP is a control group and Group SEFP and Group SERP are both experimental groups.

At first, participants completed the pretest. Then students from different groups were asked to engage in different computer lessons. The students in Group R used Interface One to read the text for 30 minutes. Students in Group SENP used Interface Two to read the text and self-explain without any prompts for 30 minutes. Students in Group SEFP used Interface Three to read the text and self-explain for 30 minutes while the system provided content-free prompts after 20 minutes. Students in Group SERP used Interface Three to read the text and self-explain for 30 minutes while the system provided content-related prompts after 20 minutes. After learning activities, participants completed a posttest.

The students in Group SEFP and Group SERP both used Interface Three. The system provided 10 prompts to students after 20 minutes. The difference was that the prompts for Group SEFP were content-free prompts while the prompts for Group SERP were content-related prompts. The content-free prompts and content-related prompts were designed in advance and both related to the same specific sentences of the text (field 3, 7, 10, 11, 12, 13, 15, 16, 18, and 21) so that the effects of different prompts can be observed. The prompts were provided to the students without
understanding the content of the students’ self-explanations. The content-free prompts referred to prompts without any domain-related or content-related information and thus could be easily adopted in any domain or content, such as “Could you explain it more clearly?”. The content-related prompts involved some domain-related or content-related information, such as “Is this a red-black tree? Why?” and “Why is node 2 located at the right child-node of node 1?”. 

Results and analyses

The results of the experiment were reported and analyzed from the following perspectives: learning effects, self-explanation generation and modification, and prompted vs. unprompted locations.

Learning effects

Among 75 participants, 11 students answered more than two questions correctly out of three questions in the RTDT pretest and thus their data were excluded. In addition, a student in Group SEFP and a student in Group SERP did not generate any self-explanations and thus their data were excluded. The assessment results of the four groups are listed in Table 1. Performance equals the mean score divided by the full marks. The pretest and posttest results of BTDT and RTDT across the four groups were not significantly different. However, the students in all four groups significantly performed better in the posttest than the pretest, both in BTDT and RTDT. This means students in the four groups gained conceptual knowledge of binary trees and red-black trees. In the Retention Test, the students in the four groups performed similarly. In the Near Transfer Test, the students in Group SEFP and Group SERP significantly performed better than the students in Group R (p < 0.05). The students in Group SENP also performed better than the students in Group R, although the difference did not reach a level of significance (p = 0.12). In the Far Transfer Test, the students in the four groups all attained low scores and thus it was possible that the floor effect might have happened. However, the students in Group SEFP and Group SERP slightly (but not significantly) performed better than the students in Group SENP and Group R. Overall, the results revealed that self-explaining through typing, particularly when prompted, made students perform better in applying target procedural knowledge to similar problems.

Table 1. Learning effects of different groups

<table>
<thead>
<tr>
<th></th>
<th>Group R n=14 mean (SD) performance</th>
<th>Group SENP n=18</th>
<th>Group SEFP n=13</th>
<th>Group SERP n=17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest (BTDT)</td>
<td>1.28 (1.14) 42%</td>
<td>1.83 (0.85) 61%</td>
<td>1.53 (0.43) 51%</td>
<td>1.47 (0.38) 49%</td>
</tr>
<tr>
<td>Pretest (RTDT)</td>
<td>0.5 (0.26) 17%</td>
<td>0.33 (0.23) 11%</td>
<td>0.30 (0.23) 10%</td>
<td>0.11 (0.11) 4%</td>
</tr>
<tr>
<td>Posttest (BTDT)</td>
<td>2.14 (0.74) 71%</td>
<td>2.55 (0.37) 85%</td>
<td>2.15 (0.47) 72%</td>
<td>2.05 (0.68) 68%</td>
</tr>
<tr>
<td>Posttest (RTDT)</td>
<td>2.42 (0.57) 81%</td>
<td>2.27 (0.56) 76%</td>
<td>2.15 (0.64) 72%</td>
<td>2.17 (1.02) 72%</td>
</tr>
<tr>
<td>Posttest - pretest (BTDT)</td>
<td>0.85 (1.51) 28%</td>
<td>0.72 (1.27) 24%</td>
<td>0.61 (0.42) 20%</td>
<td>0.58 (1.007) 19%</td>
</tr>
<tr>
<td>Posttest - pretest (RTDT)</td>
<td>1.92 (0.84) 64%</td>
<td>1.94 (0.99) 65%</td>
<td>1.84 (1.14) 61%</td>
<td>2.05 (1.05) 68%</td>
</tr>
<tr>
<td>Posttest (Retention Test)</td>
<td>0.68 (0.09) 68%</td>
<td>0.63 (0.15) 63%</td>
<td>0.8 (0.06) 80%</td>
<td>0.74 (0.10) 74%</td>
</tr>
<tr>
<td>Posttest (Near Transfer Test)</td>
<td>0.29 (0.10) 29%</td>
<td>0.49 (0.14) 49%</td>
<td>0.64 (0.06) 64%</td>
<td>0.54 (0.11) 54%</td>
</tr>
<tr>
<td>Posttest (Far Transfer Test)</td>
<td>0.02 (0.006) 2%</td>
<td>0.03 (0.004) 3%</td>
<td>0.14 (0.07) 14%</td>
<td>0.17 (0.09) 17%</td>
</tr>
</tbody>
</table>

A, B: Significantly greater than Group R, p < 0.05
C: Greater than Group R, p =0.12
Self-explanation generation and modification

The students’ self-explanations were classified according to the classifications of Chi (2000) and McNamara (2004). A self-explanation was classified as a paraphrase if the student repeated the text sentence or expressed the content of text in his/her words without further information or inference. Paraphrases were also classified into correct and incorrect paraphrases according to whether the self-explanation involved incorrect information. A classification of bridging inference was used to denote a self-explanation that integrates information across different text sentences. Prior-knowledge inference was used to indicate a self-explanation that integrates information of the text with prior knowledge. The concept of binary trees is prior knowledge of the original text, but we added additional information about the concept of binary trees in first sentence of the text in case the students did not know what a binary tree is. Therefore, self-explanations that integrate a text sentence with knowledge of binary trees were classified as bridging inferences and there was no prior-knowledge inference classified in this study. A self-explanation was classified as a logic inference if the student inferred further information from the sentence by logical deduction. Positive self-monitoring was the classification used to denote a self-explanation in which the student expressed positive understanding of the sentence, such as “It is easy”, or “I see”. Negative self-monitoring was used to indicate a self-explanation in which the student was uncertain of or questioned the text, such as “I do not know” or “Why is the node 1 black?”. Self-explanations that were beyond the text were classified as others. Table 2 lists the generating percentage of different self-explanation classifications of the final self-explanations of the different groups. The generating percentage denotes the frequency of the classifications divided by the number of text sentences, that is, 21. Inferences and self-monitoring were counted as HSE while paraphrase and others were counted as LSE. The results revealed that the students of Group SEFP and Group SERP significantly generated less incorrect logic inferences than the students of Group SENP did. It might indicate that both content-free and content-related prompts made students generate less incorrect logic inferences. In addition, the students of Group SERP generated more negative self-monitoring than the students of Group SEFP did ($p = 0.06$). It might denote that the content-related prompts made students more aware of their ignorance of the text than content-free prompts did.

Table 2. Self-explanation classifications of different groups

<table>
<thead>
<tr>
<th>Classifications/Grouping</th>
<th>Group SENP</th>
<th>Group SEFP</th>
<th>Group SERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct paraphrases</td>
<td>25.93%</td>
<td>29.3%</td>
<td>31.65%</td>
</tr>
<tr>
<td>Incorrect paraphrases</td>
<td>5.03%</td>
<td>3.66%</td>
<td>1.68%</td>
</tr>
<tr>
<td>Correct bridging inferences</td>
<td>25.66%</td>
<td>35.9%</td>
<td>33.05%</td>
</tr>
<tr>
<td>Incorrect bridging inferences</td>
<td>4.50%</td>
<td>0.00%</td>
<td>0.28%</td>
</tr>
<tr>
<td>Correct prior-knowledge inferences</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Incorrect prior-knowledge inferences</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Correct logic inferences</td>
<td>1.59%</td>
<td>1.83%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Incorrect logic inferences</td>
<td>3.97%</td>
<td>0.37%</td>
<td>1.12%</td>
</tr>
<tr>
<td>Positive self-monitoring</td>
<td>8.47%</td>
<td>0.73%</td>
<td>2.24%</td>
</tr>
<tr>
<td>Negative self-monitoring</td>
<td>8.99%</td>
<td>4.03%</td>
<td>12.88%</td>
</tr>
<tr>
<td>Others</td>
<td>2.12%</td>
<td>1.47%</td>
<td>0.84%</td>
</tr>
<tr>
<td>HSE</td>
<td>53.17%</td>
<td>42.86%</td>
<td>52.38%</td>
</tr>
<tr>
<td>HSE / LSE+HSE</td>
<td>63.19%</td>
<td>55.45%</td>
<td>60.52%</td>
</tr>
</tbody>
</table>

A, B: Significantly less than Group SENP, $p < 0.05$
C: Greater than Group SEFP, $p = 0.06$

Analyzing the records of the students’ self-explanations, the frequency and modifications of self-explanations of different groups are listed in Table 3. Self-explanation frequency indicates the total number of students’ self-explanations, including both LSE and HSE. The generating percentage indicates the frequency divided by 21, that is, the total number of text fields in the text. Modifications mean the number of students’ self-explanations that have been modified after being typed. The frequencies of self-explanations and HSE of Group SENP, Group SEFP and Group SERP are similar, but the number of modifications of Group SEFP and Group SERP are significantly greater than that of Group SENP. In addition, the number of modifications of Group SERP is significantly greater than that
of Group SEFP. The results might reveal that prompts, particularly content-related prompts, promoted students to modify their self-explanations.

Table 3. Self-explanation statistics of different groups

<table>
<thead>
<tr>
<th>Classification</th>
<th>Group SENP</th>
<th>Group SEFP</th>
<th>Group SERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-explanation frequency (SD and percentage before prompts)</td>
<td>12.61 (40.58), 60.07%</td>
<td>13.05 (42.17), 62.16%</td>
<td></td>
</tr>
<tr>
<td>HSE frequency and percentage before prompts</td>
<td>6.3 (20.2), 30%</td>
<td>6.29 (13.2), 29.9%</td>
<td></td>
</tr>
<tr>
<td>Self-explanation modifications before prompts</td>
<td>3.30 (13.23), 16%</td>
<td>3.05 (5.80), 15%</td>
<td></td>
</tr>
<tr>
<td>Final self-explanation frequency and percentage</td>
<td>18.1 (25.9), 86.24%</td>
<td>16.23 (15), 77.29%</td>
<td></td>
</tr>
<tr>
<td>Final HSE frequency and percentage</td>
<td>11.16 (36), 53.17%</td>
<td>9 (13.16), 42.86%</td>
<td></td>
</tr>
<tr>
<td>Final self-explanation modifications</td>
<td>2.44 (11.08), 12%</td>
<td>8.4 (14.89), 40%</td>
<td></td>
</tr>
</tbody>
</table>

A, B: Significantly greater than Group SENP, p < 0.05
C: Significantly greater than Group SEFP, p < 0.05

According to the classifications, self-explanation on each text field could be categorized into none (no self-explanation), LSE, or HSE, and thus the modifications of the self-explanations on the same text fields before and after prompting were classified into six kinds (Table 4). The percentage represents the frequency of the modification classification divided by the total frequency of the six classifications. The results showed that content-related prompts significantly gave rise to more self-explanation modifications, which ranged from LSE to HSE and from HSE to HSE, than content-free prompts did.

Table 4. Self-explanation modification after prompting

<table>
<thead>
<tr>
<th>Classifications/Grouping</th>
<th>Group SEFP</th>
<th>Group SERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSE to LSE</td>
<td>14.71%</td>
<td>4.88%</td>
</tr>
<tr>
<td>LSE to HSE</td>
<td>5.88%</td>
<td>18.70%</td>
</tr>
<tr>
<td>HSE to HSE</td>
<td>8.82%</td>
<td>21.14%</td>
</tr>
<tr>
<td>HSE to LSE</td>
<td>2.94%</td>
<td>1.63%</td>
</tr>
<tr>
<td>None to HSE</td>
<td>50.00%</td>
<td>39.02%</td>
</tr>
<tr>
<td>None to LSE</td>
<td>17.65%</td>
<td>14.63%</td>
</tr>
</tbody>
</table>

A, B: Significantly greater than Group SEFP, p < 0.05

Prompted vs. unprompted locations

Prompts were provided at the same specific locations in Group SEFP and Group SERP. Table 5 lists the self-explanation statistics on prompted locations (10 fields) and un-prompted locations (11 fields) of different groups. The percentage represents the frequency divided by the total number of fields. The results showed that the percentages of HSE and self-explanation modifications in prompted locations were significantly higher than that in un-prompted locations both in Group SEFP and Group SERP. The results of comparing different groups also showed that the percentage of self-explanation modifications on prompted locations in Group SEFP and Group SERP was both significantly greater than that in Group SENP. It revealed that both content-free and content-related prompts promoted more self-explanation modifications in prompted locations. In addition, the percentage of self-explanation modifications in un-prompted locations in Group SERP was also greater than that in Group SENP. It might indicate that content-related prompts also promoted students to generate more self-explanation modifications even in unprompted locations. However, the results also showed that the percentage of HSE in un-prompted locations in Group
SENP was significantly greater than that in Group SEFP and Group SERP. It might indicate that prompts made students pay more attention in prompted locations and less attention in un-prompted locations.

| Table 5. Self-explanation statistics on prompted and un-prompted locations |
|-------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                                  | Group SENP      |                  | Group SEFP      |                  | Group SERP      |
|                                                  | Prompted locations | Un-prompted locations | Prompted locations | Un-prompted locations | Prompted locations | Un-prompted locations |
| Self-explanations                               | 8.44 (8.26) 84% | 9.66 (5.2)  87% | 8.6 (1.75)  86% | 7.53 (10.7)  68% | 8.64 (5.1)  86% | 8.76 (6.06)  79% |
| HSE                                             | 5.77 (9.7)  57% | 5.38 (10.4) 48.9% | 5.5 (5.6)  55% | 3.3 (2.89)  30% | 7 (5.8)  70% | 3.17 (5.4)  28% |
| Self-explanation modifications                   | 1.05 (4.4) 10% | 1.44 (3)  13% | 5.15 (4.8) 51% | 2.5 (5.6)  23% | 7.29 (8.9) 72% | 3.17 (7) 28% |

A: Approximately significantly greater than un-prompted locations, $p = 0.053$
B, C, D, E: Significantly greater than un-prompted locations, $p < 0.05$

a, f: Significantly greater than Group SEFP, $p < 0.05$
b: Approximately significantly greater than Group SEFP, $p = 0.053$
c: Significantly greater than Group SERP, $p < 0.05$
d, e, g: Significantly greater than Group SENP, $p < 0.05$

Discussion

This study focuses on content-free computer supports because these supports can be easily applied to other domains or content. The reading and typing self-explaining interface can be applied to other domains or content by changing the text, but the typing interface makes students’ self-explanations more complex to analyze than a menu-based interface because the analysis involves natural language processing. This study also used a learning companion to provide prompts without understanding the content of students’ self-explanations. This releases the requirement to build domain-related knowledge into the system to analyze the self-explanations. Content-free prompts can be used for any domain or content, and content-related prompts can be assigned in advance along with the text. Thus the developed system can be used for other domains and content by changing the text and content-related prompts.

The prompting mechanism in this study remains several issues unclear and some room for enhancement. First, this study provided prompts without understanding the content of students’ self-explanations. On the other hand, understanding the students’ self-explanations can enable system to provide students with adaptive prompts and may have better effects on self-explanations. However, whether adaptive prompts could promote better effects remains unclear. Second, this study provided prompts after 20 minutes for all students. Prompting timing could be changed to fit the reading and self-explaining speed of different students. For example, the students push a button when they complete self-explaining on the text and then the system provides prompts. Thirdly, this study prompted at the same specific locations for all students. The prompting number and locations could be randomly assigned or be adaptive, such as prompting at the sentences where the students did not generate any self-explanations.

Some qualitative findings are reported below, but most of them are preliminary observations and require further investigation.

The study of Hausmann and Chi (2002) found that typed self-explanations can be noted for their completeness and spoken self-explanations tend to be fragmented and incoherent. They also pointed out that “in typing, students might have filtered out what they would spontaneously say orally.” Consistent with their finding, students’ self-explanations in this study appear to be complete. However, some students (about 23%) typed some self-explanations to express what they were thinking, such as “hmm…”, “Oh!”, “I am still thinking …”. These students might tend to be less filtered when typing their self-explanations.
The interface used in this study allowed students to read and self-explain in their own sequences. Analyzing the students’ reading and self-explaining sequences, three kinds of sequences were recognized. Some students (about 56.25%) read several text fields forward and backward many times before typing self-explanations on a field. As an example, most of them typed self-explanations on field 1 after reading field 1 to field 3 or field 5. Some students (about 15%) read all text fields once and then began to self-explain from field 1. Some students (about 15%) read and self-explained almost field by field and they seldom read fields forward and backward; that is, they read a field, self-explained the field, and then read the next field. Other students generated few self-explanations and their sequences were not recognized. The above results reveal that students have different reading and self-explaining sequences. These sequences might be regarded as records of students’ effort to understand the text. However, many issues remain unclear, for example, “Do these sequences matter?” and “Can these sequences help us understand their self-explanations?”

The interface used in this study allowed students to modify their previous self-explanations. The results in Table 3 reveal that students will modify their previous self-explanations. Their modifications were analyzed to clarify how they modified their self-explanations. Three kinds of modifications were classified: addition, deletion, and adjustment. First, students added more self-explanations on some fields in addition to previous self-explanations. The self-explanations that were added may be paraphrases, inferences, or self-monitoring. Most modifications in this study were additions. Secondly, students deleted some self-explanations from their previous self-explanations. This might indicate that students found these self-explanations incorrect and thus deleted them. Thirdly, students adjusted the content of their previous self-explanations by adding, deleting, or changing some words. The adjustments tended to be minor and to make the self-explanations more complete. Different kinds of modifications may occur simultaneously. For example, a student deleted a self-explanation from his previous self-explanations and then added some self-explanations as a modification. However, these modifications of self-explanations might be regarded as symbols that students revised their understanding on the text. It could also justify our assumption that a typing record might provide an opportunity for students to reflect on or revise their understanding of the text.

Comparing to the few HSE generating percentages (1.6% without prompts and 10.3% with prompts) in the study of Hausmann and Chi (2002), the students in the full-text reading and typing self-explaining environment generated higher HSE generating percentages (about 43%–53% on different groups). There might be several possible reasons to account for the difference between these two results. First, the full-text reading and modifiable typing self-explaining interface allows students to modify their previous self-explanations and to reflect on their previous understanding of the text and thus might provide a more constructive environment for students. Secondly, participants in this study major in computer science and may be used to typing. However, their attitudes toward typing and typing speeds were not evaluated. Thirdly, the study of Hausmann and Chi (2002) adopted a text-only situation to present the text, and this study adopted a text situation in the first part of the text and a diagram situation in the second part. Roy and Chi (2005) suggested that text in diagram or multimedia situations promotes higher percentages and performance of students’ self-explanations than text in a text-only situation does. On the other hand, Roy and Chi (2005) listed the proportions of HSE to the sum of HSE and LSE of some spoken self-explaining studies in different learning contexts: average 45.04% in a text situation, average 68.26% in a multimedia situation, and 91.61% in a diagram situation. In addition, the proportion in the typing self-explaining study of Hausmann and Chi (2002) is 8.45%. In this study, the proportions of HSE to the sum of HSE and LSE are 55%–63% on different groups. Although many variables in this study and the study of Hausmann and Chi (2002) were not equivalent (for example, participants and materials), and thus many issues require more controlled experiments for clarification, this study reveals that self-explaining through typing could make similar high HSE proportion as spoken self-explaining.

Conclusion

This study investigated the effects of two content-free computer supports for self-explaining: modifiable typing interface and prompting. The full-text reading and modifiable typing interface allowed students to read text in their own sequences, type self-explanations, and modify their previous self-explanations. In addition, the computer could provide prompts to promote students’ self-explanations. The results of the experiment showed that the students that self-explained through typing, particularly those being prompted, performed better in applying target procedural knowledge to similar problems than the reading students did. The results also revealed that typing record of self-explanations might provide an opportunity for students to reflect on or revise their understanding of the text.
This study also investigated the effects of prompting by a learning companion without understanding the content of students’ self-explanations. Comparing content-free and content-related prompts, some similar effects existed. First, the results showed that both content-free and content-related prompts made self-explaining students performed better in applying target procedural knowledge to similar problems than the reading students did. Secondly, they both promoted students to generate more self-explanation modifications and less incorrect logic inferences. Thirdly, they both made students generate higher HSE percentage and self-explanation modifications in prompted locations than un-prompted locations. It might indicate that prompts made students pay more attention in prompted locations and less attention in un-prompted locations.

However, content-related prompts facilitated more self-explanation modifications and more negative self-monitoring than content-free prompts did. Analyzing the classifications of self-explanation modifications, content-related prompts promoted more modifications, which ranged from LSE to HSE and from HSE to HSE, than content-free prompts did. The possible cause for the above results might be that content-related prompts asked the students about specific questions about the text and promoted students to reflect on their understanding or be aware of their ignorance.

Acknowledgments

The authors would like to thank the support of the National Science Council (NSC94-2520-S-155-002). The authors also thank two anonymous reviewers for their valuable comments to revise the paper.

References


Examining the Factors Influencing Participants' Knowledge Sharing Behavior in Virtual Learning Communities

Irene Y.L. Chen¹, Nian-Shing Chen² and Kinshuk³

¹Department of Information Management, Ching Yun University, Taiwan // irene@cyu.edu.tw
²Department of Information Management, National Sun Yat-sen University, Taiwan // nschen@faculty.nsysu.edu.tw
³School of Computing and Information Systems, Athabasca University, Canada // kinshuk@ieee.org

ABSTRACT

Increasing organizations and educational institutions have implemented virtual learning communities to encourage knowledge sharing. However, this task can not be accomplished simply by grouping people together and telling them “sharing your knowledge will make you learn better”. This research attempts to examine the factors influencing knowledge sharing from the perspective of human behavior. Theory of Planned Behavior is integrated with social network ties and empirical findings from virtual learning community literature to develop the research model. The current research model comprises eight hypotheses to explore questions of whether social network ties, learners’ attitude toward knowledge sharing, learners’ beliefs of their capabilities in performing online knowledge sharing, and subjective norms relate to knowledge sharing intention, which leads to actual behavior in a virtual learning environment. This study empirically validates the hypothesized relationships using a field survey of college students and MBA students enrolling courses conducted in a virtual learning community. Attitude, subjective norm, Web-specific self-efficacy and social network ties are shown to be good predictors of knowledge sharing intention which, in turn, is significantly associated with knowledge sharing behavior. Knowledge creation self-efficacy does not significantly impact knowledge sharing intention.

Keywords

Virtual learning community, Theory of planned behavior, Social network

Introduction

The explosion in Web based technology has led to increasing volume and complexity of knowledge, which stimulates the proliferation of virtual learning communities (VLCs). They are information technology based cyberspaces in which individuals and groups of geographically dispersed learners accomplish their e-learning goals (Yang et al., 2007). Such kind of communities has become growing initiatives during the past years in business organizations, educational institutions, and governments to pursue and mobilize knowledge via the Internet.

Despite the proliferation of virtual learning communities in modern life, very little is known about factors leading to their success or failure. The objective of VLCs is to enhance learning performance by encouraging participants to exploit or explore knowledge. Therefore, whether learners can share knowledge by raising good questions, recommending good articles, providing ideas and helping others resolve problems over Internet become critical in enhancing their learning performance. As Leonard (1995) indicated, valuable knowledge collects in individuals’ head and is embodied in machines, software, and routine processes. The participants of learning activities in VLCs need to understand precisely what knowledge will fulfill their needs, and to keep this knowledge on the cutting edge, deploy it, leverage it in performing their tasks, and spread it across the community.

While knowledge sharing has been found to be a motivation for using emergent virtual communities (VC) (Wasko & Faraj, 2000), a range of barriers has been found impeding online knowledge sharing. The barriers having been discussed in VC literature include poor quality of community websites, members’ lack of social relationship ties in the social network, their lack of capabilities in using community websites and involving in knowledge sharing, and the cognitive process such as negative attitude toward sharing knowledge online (Chen, 2007; Malhotra & Galletta, 2004; Sangwan, 2005). These barriers, however, have received little attention in VLC literature. Although some recent VLC studies have explored the importance of individuals’ attitude or self-efficacy in performing online knowledge sharing, their results only demonstrated the roles of an individual’s belief and capabilities in performing such behavior, without considering the effects of an individual’s centrality in a social network and the opinions provided by others in such network (subjective norm). Wasko and Faraj (2005) pointed out that an individual’s structural position in an electronic network of practice will influence his or her willingness to contribute knowledge to others. Given the social characteristic of virtual learning communities, this study suggests that attitude, self-
efficacy, social network ties and subjective norm will collectively shape knowledge sharing behavior. The Theory of Planned Behavior (TPB) is used as the underlying framework for developing the research model.

TPB has been one of the most influential theories in explaining and predicting behavior, and it has been shown to predict a wide range of behaviors (Sheppard et al., 1988). TPB posits that individuals’ behavior is determined by behavioral intention and perceived behavioral control. Behavioral intention is determined by attitude toward behavior (ATT), subjective norm (SN), and perceived behavioral control (PBC) (Ajzen, 1985). Attitude toward behavior reflects one’s favorable/unfavorable feelings of performing a behavior. Subjective norm reflects one’s perception of relevant others’ opinions on whether or not he or she should perform a particular behavior. Perceived behavioral control reflects one’s perceptions of the availability of resources or opportunities necessary for performing a behavior (Ajzen & Madden, 1986). Recently, some academicians found that TPB can be used as a theoretical guidance for explaining knowledge sharing intention (e.g., Bock et al., 2005) or online behavior. In investigating online consumer behavior, Pavlou and Fygenson (2006) further suggested that the nature of PBC comprises two distinct dimensions: self-efficacy (SE) and controllability. SE is defined as individual judgments of a person’s capabilities to perform a behavior. Controllability is defined as individual judgments about the availability of resources and opportunities to perform the behavior. Bandura (1997) suggested that self-efficacy is one of the major cognitive forces guiding behavior.

Despite an emerging interest among IS researchers toward the knowledge sharing phenomenon, there is only a limited and fragmented understanding of online knowledge sharing behavior in virtual learning communities. Although TPB has been used to explain human behavior in different situation, it can not fully explain the online knowledge sharing behavior in the VLC setting. The three characteristics of a VC are as follows (Weigand et al., 2000). First, it consists of a group of people with shared interests. Second, the role of the enabling (Internet) technologies is more than just passively capturing and representing information. They should also help unify the community by actively engaging users in defining and integrating the information resources of the community. Third, the members of the community should not be forced to use the technologies. The participants in most VLCs include teachers and learners. These people share a common goal of learning by sharing knowledge over Internet. Learners’ capabilities in using the website functions to acquire and provide knowledge are essential prerequisites for performing online knowledge sharing in such environment. The lack of confidence in using the website functions may diminish learners’ intention to perform online knowledge sharing.

On the other hand, to encourage learners to use the VLC website as a communication channel, meaningful content must be continuously created to enrich the resource of the VLC’s electronic knowledge repository. Learners’ capabilities in creating useful knowledge to others are also important in a VLC where people are expected to learn anytime anywhere. According to social cognitive theory, one’s belief about his or her capabilities in performing a particular behavior is defined as self-efficacy (Bandura, 1997). It has been found to vary across activities and situational circumstance. Prior literature has discussed several types of self-efficacy related to learning performance, such as mathematics self-efficacy (Banta, 1989; Pajares, 1996), computer self-efficacy (Compeau and Higgins, 1999; Marakas et al., 1998), etc. Marakas et al. (1998) and Agarwal et al. (2000) both suggested that a distinction be made between general computer self-efficacy (an individual’s judgment of efficacy across multiple computer application domains) and software-specific self-efficacy (individual’s feeling of self-efficacy relative to a specific software package). This study hence focuses on web-specific self-efficacy (WBSE) and knowledge creation self-efficacy (KCSE) for the current research purpose. WBSE refers to a learner’s beliefs about his or her capabilities in using the functions of the VLC website. KCSE refers to a learner’s beliefs about his or her capabilities in articulating the ideas and experiences, synthesizing knowledge from different sources, and learning from others by embodying explicit knowledge into tacit knowledge.

Moreover, knowledge experiences network effects as its value increases as more people share it. Online knowledge sharing behavior in VLCs has some notable differences compared to face-to-face group discussion conducted in traditional classrooms. The spatial and temporal separation between learners and teachers increases difficulties of effective ideas exchange. Therefore, the ties that connect individuals in a network are required for the sharing of knowledge. The concept of social network ties is derived from social capital theory (SCT) which has been widely applied in social science disciplines. In SCT-based literature, social network ties are defined as channels for information and resource flows (Tsai & Ghoshal, 1998). Yang et al. (2007) utilized the concept of social network ties to develop a peer-ranking mechanism that is useful for helping learners find quality learning content and trustworthy learning collaborators. Wiberg (2007) suggested the concept of learning through networks as a challenging concept.
for addressing user-driven technologies that support social, collaborative and creative learning processes in, via, or outside typical educational settings. Tomsic and Suthers (2006) suggested that the formation of new collaborative ties is more significant for learning through information sharing in social networks than raw frequency of interaction. The ties that link the members of a social network like a VLC can be helpful for encouraging individuals to engage in knowledge sharing activities.

This study hence aims to better explain online knowledge sharing behavior in VLCs by extending the TPB with social network ties and dividing the self-efficacy into two sub-dimensions: learners’ knowledge creation self-efficacy and learners’ web specific self-efficacy. The proposed research model posits that attitude toward online knowledge sharing, subjective norm, learners’ knowledge creation self-efficacy, learners’ web-specific self-efficacy, and social network ties jointly determine learners’ knowledge sharing intention which, in conjunction with learners’ knowledge creation self-efficacy and learners’ web-specific self-efficacy determines knowledge sharing behavior.

**Literature Review and Research Model**

**Virtual Learning Community**

The rise of the electronic-learning (e-learning) has been considered to be an important stimulus for the interest in virtual learning communities (VLCs) during the last decade. A VLC is the information technology based cyberspace in which individuals and groups of geographically dispersed learners accomplish their e-learning goals. It is also a place with socially constituted values and expectations (Burnett & Dickey, 2003). One of the VLCs’ purposes is to encourage knowledge sharing so that valuable knowledge embedded in the network can be effectively explored. Most of the learners participate in VLCs with the expectations that they can acquire and share valuable knowledge to fulfill their needs.

Consistent with the growing interest in online cooperation and virtual organizing, there has been an exponential growth of VLC studies. For example, Daniel et al. (2003) addressed the importance of social capital in enhancing learning performance in virtual learning communities. Haythornthwaite (1998) found that centrality in a network may correlate with performance or satisfaction measures. Piccoli et al. (2001) found that students in the virtual learning environment will report higher levels of computer self-efficacy than their counterparts in the traditional learning environment. Markland (2003) pointed out that the 'new alliances' approach is most effective at individual as well as at group level for information resource seeking in virtual learning communities. Wellman et al. (1996) asserted that computer-supported social networks can maintain strong, supportive ties as well as increase the number and diversity of weak ties in virtual communities. These ties make on-line communications more uninhibited and creative. Wasko and Faraj (2005) investigated why people share their knowledge with others in electronic network of practice. They found that both reputation and centrality have significant influences on the helpfulness and volume of knowledge contribution. Finally, Koh and Kim (2004) investigated the relationship between community knowledge sharing activity and the loyalty toward the virtual community service provider. The results indicated that the level of community knowledge sharing activity may be a proper proxy for the state of health of a virtual community.

Prior studies have provided evidence demonstrating the importance of knowledge exchange in enhancing the learning performance. While knowledge sharing has been found to be a motivation for using emergent VCs (Wasko & Faraj, 2000), little research has examined what is required for effective knowledge sharing in VLCs. Davenport and Prusak (1998) indicated that sharing knowledge is often unnatural; hoarding knowledge and looking suspiciously upon knowledge from others are the natural tendency. Most people are reluctant to share their knowledge due to the fears of losing superiority. This human nature necessitates a study to examine the human-related factors of online knowledge sharing. In recent years, considerable studies have elaborated on the investigation of the determinants of individual’s knowledge sharing behavior in the communities of practice or virtual communities. Virtual learning community is, however, one specific type of VC and has so far received little attention on what can influence participants’ knowledge sharing behavior. Based on a review of virtual learning community and information system (IS) literatures, this study integrates the concept of social network ties into the Theory of Planned Behavior to provide a model which aims to better explain substantial portion of variance of such behavior in the VLC context.
To develop the research model, the VLC and IS literatures were reviewed to identify factors that potentially affect knowledge sharing in virtual learning communities. This section elaborates on the theory base and derives the hypotheses. The research model is depicted in Figure 1.

**Intentions and Actual Behavior of Knowledge Sharing**

Behavioral intention has long been found to be significantly associated with actual behavior. According to the theory of planned behavior (Ajzen, 1991), behavioral intentions are motivational factors that capture how hard people are willing to try to perform a behavior. TPB suggests that behavioral intention is the most influential predictor of behavior; after all, a person does what she intends to do (Pavlou & Fygenson, 2006).

Prior literature has corroborated the relationship between the two variables. For example, in a meta-analysis of 87 studies, an average correlation of .53 was reported between intentions and behavior (Sheppard et al., 1988). Results of Pavlou and Fygenson’s (2006) longitudinal study validated the predictive power of TPB in online behavior and showed strong associations between get-information intention and get-information behavior, and between purchase intention and purchase behavior. Also, the Theory of Implementation Intentions (Gollwitzer, 1999) holds that a goal-driven behavior automatically activates a set of goal-enabling (implementation) intentions that help realize the behavior (Sheeran & Orbell, 1999). In this study, receiving good grades is learners’ common goal and sharing knowledge online is a goal-driven behavior which can be realized if learners intend to perform such behavior. A positive relationship between intentions to share knowledge online and its actual behavior is thus expected.

**H1:** Knowledge sharing intention positively influences online knowledge sharing behavior.

**Subjective Norm**

Subjective norm suggests that behavior is instigated by one’s desire to act as important referent others act or think one should act (Pavlou & Fygenson, 2006). Applied to the focal behavior, SN reflects participant perceptions of whether the behavior is accepted, encouraged, and implemented by the participant’s circle of influence. The literature suggests a positive relationship between SN and intended behavior, and empirical work has shown that SN influences behavioral intentions toward system use (Karahanna et al., 1999). SN has been shown to be an important determinant of acceptance behaviors in numerous studies (Karahanna & Straub, 1999; Srite & Karahanna, 2006; Taylor & Todd, 1995; Thompson et al., 1991; Venkatesh & Davis, 2000; Venkatesh & Morris, 2000; Venkatesh et
al., 2003). Bock et al. (2005) conducted a survey with thirty organizations to test a knowledge sharing model. Results suggested that subjective norm has significant influence on knowledge sharing intention. Raaij and Schepers (2006) conducted a survey with 45 Chinese participants in an executive MBA program. They found that subjective norm has indirect effects on the use of virtual learning communities. One’s social environment is a valuable source of information to reduce uncertainty and determine whether behaviors are within rules and are acceptable. Therefore, subjective norms may, through informational and normative influences, reduce uncertainty with respect to whether use of a system is appropriate (Evaristo & Karahanna, 1998; Srite & Karahanna, 2006). A positive relationship between SN and intentions to share knowledge online is thus expected.

H2: Subjective norm positively influences knowledge sharing intention.

Attitude

Attitude has long been shown to influence behavioral intentions (Ajzen & Fishbein, 1980). This relationship has received substantial empirical support (Pavlou & Fygenson, 2006). For example, Bock et al. (2005) conducted a survey with thirty organizations to test a knowledge sharing model. Results suggested that attitude toward knowledge sharing positively and significantly influence behavioral intention. Brown and Venkatesh (2005) employed the TPB framework to propose a model presenting factors influencing household technology adoption. They concluded that attitude toward IT usage positively influences technology adoption intention. Galletta et al. (2006) investigated the interacting effects of website delay, familiarity and breadth on users’ performance of information search. They found a positive and significant relationship between attitude and behavioral intention. Kolekofski and Heminger (2003) proposed a model that defines the influences on one’s intention to share information. Using the workers in a unit of a large governmental organization as samples, their study confirmed that attitude influences intention to share information. Finally, the integrated model proposed by Wixom and Todd (2005) empirically validated the strong association between attitude toward IS use and intention to use IS.

Consistent with these studies, Bhattacherjee and Premkumar (2004) found that attitude is important to predict the intention of continuing use of the computer-based training system usage and rapid application development software; attitude is one of the key perceptions driving users’ IT use behavior. Agarwal and Prasad (1999) expanded on their prior studies, which utilized learning theory and various individual difference variables, to address possible causes for inconclusive results of prior studies. They found that attitude is an important antecedent of individuals’ behavioral intention. Following the TPB literature, a positive relationship between attitude and intentions to share knowledge online is expected.

H3: Participants’ attitude toward online knowledge sharing positively influences knowledge sharing intention.

Perceived Behavioral Control

In this study, perceived behavioral control refers to the learner’s perceived ease or difficulty of telling story and experiences, writing documents, expressing opinion. Some researchers have suggested that PBC composes two distinct dimensions: self-efficacy (SE) and controllability (e.g., Pavlou & Fygenson, 2006). Pavlou and Fygenson (2006) defined self-efficacy as individual judgments of a person’s capabilities to perform a behavior, and controllability as individual judgments about the availability of resources and opportunities to perform the behavior. Given that each participant has equal opportunities to access the anytime, anywhere available resource in a VLC, this study focuses on self-efficacy to discuss its role in the formation of learners’ decision to perform knowledge sharing.

Self-efficacy has been found to vary across activities and situational circumstance. For example, Joo et al. (2000) found that Internet self-efficacy is able to predict students’ performance on search task in Web-based instruction. Thompson et al. (2002) concluded that task-specific Internet self-efficacy has a significant effect on online search performance. Marakas et al. (1998) indicated that task-specific computer self-efficacy is an individual’s perception of efficacy in performing specific computer-related tasks within the domain of general computing. Other types of self-efficacy such as social self-efficacy (Sherer et al., 1982), teaching self-efficacy (Gibson & Dembo, 1984), math self-efficacy (Banta, 1989) have been applied in various domains to explore their predictive powers on task
performance, one’s decision on what behaviors to undertake, and the level of commitment and persistence in attempting those behaviors.

**Web-Specific Self-Efficacy and Knowledge Creation Self-Efficacy**

In their TPB-based study, Pavlou and Fygenson (2006) found that self-efficacy is a good predictor of online consumer behavior. In examining electronic knowledge repositories (EKR), Kankanhalli et al. (2005) proposed knowledge self-efficacy (KSE) which can be manifested in the form of people believing that their knowledge can help to solve job-related problems, improve work efficiency, or make a difference to their organization (p122.). They found that KSE is positively related to EKR usage by knowledge contributors. Given that a VLC is an Internet-technology-based environment, when learners are required to learn in such setting, they need to be capable of using the website functions, and exploiting or exploring existing knowledge resource. Exploitation involves capturing existing knowledge and transferring and deploying knowledge in other similar situations, while exploration needs sharing and synthesizing existing knowledge, and creating new knowledge. These processes are defined by Nonaka (2000) as the methods of knowledge creation. Therefore, this study focuses on two types of learner self-efficacy: Web-specific self-efficacy (WBSE) and knowledge creation self-efficacy (KCSE).

WBSE refers to a learner’s beliefs about his or her capabilities in using the functions of the VLC website. KCSE is defined as a learner’s beliefs about his or her capabilities to articulate the ideas and experiences, synthesize knowledge from different sources, and learning from others by embodying explicit knowledge into tacit knowledge. The learners that have higher WBSE and KCSE will gain greater benefit by participating in VLC activities than those that have lower WBSE and KCSE. These cognitive beliefs affect how learners process knowledge and learn through a VLC. Thus, it seems plausible that learners’ beliefs about their capabilities in using the website functions and creating new knowledge in VLCs could influence their propensity to undertake knowledge-based activities in such environment. Positive relationships between two types of self-efficacy and intentions and behavior of online knowledge sharing are thus expected.

H4: Knowledge creation self-efficacy positively influences knowledge sharing intention.
H5: Knowledge creation self-efficacy positively influences knowledge sharing behavior.
H6: Web-specific self-efficacy positively influences knowledge sharing intention.
H7: Web-specific self-efficacy positively influences knowledge sharing behavior.

**Social Network Ties**

Virtual learning communities provide a suitable environment for those seeking for knowledge-based communication. Learners in a VLC may or may not know each other, yet they sometimes communicate actively through the cyberspace. With no immediate benefit to the knowledge contributors, and free-riders are able to acquire the same knowledge as everyone else (Wasko & Faraj, 2005, p.35), why learners exchange ideas and experience with strangers is not well understood. Some academicians (Cohen & Prusak, 2001; Kankanhalli et al., 2005; Nahapiet & Ghoshal, 1998; Wasko & Faraj, 2005) have attributed such behaviour to social network ties, an important attribute of social capital derived from the Social Capital Theory.

Social network ties are channels for information and resource flows (Tsai & Ghoshal, 1998). They can be considered as a bond between two people based on one or more relations they maintain in a social network (Haythornthwaite, 1998). Academicians have addressed the importance of social interaction ties in the creation or exchange of knowledge and mobile learning. For example, Chen et al. (2008) proposed the architecture of a mobile learning management system which can better support mobile learning for a small group of learners with effective social interaction. Tsai and Ghoshal (1998) found that social interaction ties had direct positive impacts on the extent of interunit resource exchange. Chen (2007) found that social interaction ties can enhance individuals’ intentions to perform online knowledge sharing. Finally, Yang et al. (2007) suggested that posting and responding to messages can help to create social-interaction ties among the members of a VLC, these ties are helpful for encouraging online knowledge sharing. A positive relationship between social network ties and the intentions of online knowledge sharing is hence expected.
H8: Social network ties positively influence knowledge sharing intention.

Research Methodology

Data Collection

Data for this study were collected in two phases from 396 full-time senior college students and MBA students who enrolled in two courses (enterprise resource planning and electronic business) that were conducted over a Web-based learning platform and physical classrooms. These students majored in different areas (information management, business administration and graduate school of management). Instructors upload the course material and the topic to be discussed one week prior to each class and students read the material, upload their completed homework assignments and share their opinions or ideas in the discussion forum. The students then meet in the physical classroom to receive the scored homework assignments and comments from the instructors. Generally, students in the same class knew each other; those who were in different classes but enrolled in the same course might not know each other. All of the subjects had received a basic training of WWW use prior to the survey (in the first two years of their college life).

In the first week of the semester, the instructors explained how the course activities will be conducted and how learners’ performance will be evaluated at the end of the semester. The first-phase survey questionnaires measuring learners’ attitude towards knowledge sharing and their self-efficacy were administered to learners in the second week. The second-phase survey questionnaires concerning subjective norms, social network ties, knowledge sharing intention and knowledge sharing behavior were given to learners in the last week and responses were collected in the same week.

The instrument was examined to ensure content validity and reliability within the target context. A pretest of the questionnaire was performed using 3 experts in the IS area to assess logical consistencies, ease of understanding, question item sequence adequacy, and context fitness. Necessary modifications were made according to the comments collected from these experts. Furthermore, a pilot study was conducted using another 2 PhD. students and 5 master students who had taken courses conducted in a virtual learning community. Comments and suggestions on question item content and structure were solicited and proper modification was made accordingly.

Construct measurement

Most of the scales were drawn from prevalidated measure in VC or KM literature. Specifically, knowledge sharing behavior was measured by the items adapted from Davenport and Prusak (1998) and Wasko and Faraj (2005). Intention of knowledge sharing was measured by the items adapted from Bock et al. (2005) and Kankanhalli et al. (2005). Items for measuring the attitude and subjective norm were adapted from prior related research (Bock et al., 2005; Brown & Venkatesh, 2005; Pavlou & Fygenson, 2006; Srite & Karahanna, 2006). The instruments for knowledge creation self-efficacy and web-specific self-efficacy were adapted from Kankanhalli et al. (2005) and Pavlou and Fygenson (2006). Finally, social network tie was measured by items adapted from prior related research (Nahapiet & Ghoshal, 1998; Tsai & Ghoshal, 1998; Yli-Renko et al., 2002). For the aforementioned measures, a seven point Likert type was used, with anchors ranging from 1 (strongly disagree) to 7 (strongly agree). Appendix A shows all the measurement items.

Data Analysis

Measurement Model Testing

Construct validity for the seven measurement scales with a LISREL confirmatory factor analysis (CFA) was assessed. The measurement model in the CFA was revised by dropping items, one at a time, which shared a high degree of residual variance with other items, according to the standard LISREL methodology (Gefen, 2000; Gerbing & Anderson, 1988). Items were dropped depending on reported standardized residuals, that is, those showing a significant degree of shared non-specified variance among the measurement items. Composite reliability was then
examined using the Cronbach’s alpha values. As listed in Table 1, all of these values were greater than 0.82, well above the recommended threshold value of 0.7.

<table>
<thead>
<tr>
<th>Construct and Scale Items</th>
<th>Item</th>
<th>Std. Deviation</th>
<th>Factor Loading</th>
<th>Construct Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT (ATT)</td>
<td>ATT1</td>
<td>4.44</td>
<td>0.64</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>ATT2</td>
<td>4.36</td>
<td>0.63</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>ATT3</td>
<td>4.33</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>ATT4</td>
<td>4.44</td>
<td>0.57</td>
<td>0.75</td>
</tr>
<tr>
<td>SN (SN)</td>
<td>SN1</td>
<td>4.05</td>
<td>0.76</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>SN2</td>
<td>3.98</td>
<td>0.84</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>SN3</td>
<td>4.08</td>
<td>0.74</td>
<td>0.71</td>
</tr>
<tr>
<td>KCSE (KCSE)</td>
<td>KCSE1</td>
<td>3.89</td>
<td>0.79</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>KCSE2</td>
<td>3.90</td>
<td>0.87</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>KCSE3</td>
<td>3.93</td>
<td>0.79</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>KCSE4</td>
<td>3.93</td>
<td>0.81</td>
<td>0.74</td>
</tr>
<tr>
<td>WBSE (WBSE)</td>
<td>WBSE1</td>
<td>3.93</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>WBSE2</td>
<td>3.92</td>
<td>0.74</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>WBSE3</td>
<td>3.82</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>WBSE4</td>
<td>4.10</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td>SNT (SNT)</td>
<td>SNT1</td>
<td>4.09</td>
<td>0.82</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>SNT2</td>
<td>3.99</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>SNT3</td>
<td>4.20</td>
<td>0.70</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>SNT4</td>
<td>4.23</td>
<td>0.70</td>
<td>0.72</td>
</tr>
<tr>
<td>INT (INT)</td>
<td>INT1</td>
<td>4.24</td>
<td>0.68</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>INT2</td>
<td>4.35</td>
<td>0.64</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>INT3</td>
<td>4.25</td>
<td>0.63</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>INT4</td>
<td>4.16</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>BEH (BEH)</td>
<td>BEH1</td>
<td>4.51</td>
<td>0.59</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>BEH2</td>
<td>4.35</td>
<td>0.61</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>BEH3</td>
<td>4.44</td>
<td>0.62</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>BEH4</td>
<td>4.43</td>
<td>0.62</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Discriminant validity was evaluated through comparison of the AVE of construct pairs to the squared correlation between construct pairs. Fornell and Larcker (1981) recommended a test of discriminant validity, where the AVE should exceed the squared correlation between that and any other construct. All of the items in Table 2 exhibited that all AVEs are greater than the highest squared correlation (0.481) between construct pairs. The overall results provide support for acceptable discriminant validity of constructs.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Squared Correlations and AVEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT</td>
<td>SN</td>
</tr>
<tr>
<td>ATT</td>
<td>.71</td>
</tr>
<tr>
<td>SN</td>
<td>.233(**)</td>
</tr>
<tr>
<td>KCSE</td>
<td>.142(**)</td>
</tr>
</tbody>
</table>
Structural Model Testing

For models with good fit, it is suggested that chi-square normalized by degrees of freedom ($\chi^2/df$) should not exceed 5 (Bentler & Bonett, 1980). Etezadi-Amolo and Farhoomand (1996) and Henry and Stone (1994) both suggested that the recommended thresholds of GFI and AGFI are 0.80. Chau (1996) indicated that the NNFI and CFI with values greater than 0.85 are moderate fit indices. For the current CFA model, $\chi^2/df$ was 2.54, NFI was 0.95, NNFI was 0.93, CFI was 0.95, GFI was 0.92, AGFI was 0.89. SRMSR was 0.03, slightly above the recommended value (see Table 3). These figures signify an adequate model fit.

**Table 3. Goodness of Fit Statistics**

<table>
<thead>
<tr>
<th>Construct Guideline</th>
<th>Fit Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2/df$</td>
<td>&lt;=5.0</td>
</tr>
<tr>
<td>NFI</td>
<td>&gt;=0.85</td>
</tr>
<tr>
<td>NNFI</td>
<td>&gt;=0.85</td>
</tr>
<tr>
<td>CFI</td>
<td>&gt;=0.85</td>
</tr>
<tr>
<td>GFI</td>
<td>&gt;=0.80</td>
</tr>
<tr>
<td>AGFI</td>
<td>&gt;=0.80</td>
</tr>
<tr>
<td>SRMSR</td>
<td>&lt;=0.1</td>
</tr>
</tbody>
</table>

Figure 2. Structural model test
The path significance of each hypothesized association in the research model and variance explained (R² value) by each path were examined. All the paths were significant except the one pointing from knowledge creation self-efficacy to knowledge sharing intention. Knowledge sharing behavior (BEH) was predicted by knowledge sharing intention, web-specific self-efficacy and knowledge creation self-efficacy, which jointly explained 52% of the behavior variance. Knowledge sharing intention was predicted by subjective norm, attitude, web-specific self-efficacy, knowledge creation self-efficacy and social network ties, which jointly explained 59% of the intention variance. As can be seen from Figure 2, hypothesis H4 was not supported whereas the rest of the hypotheses were all supported.

Discussion

Summary of results

The proposed research model provides a framework for understanding online knowledge sharing behavior in virtual learning communities. Learners’ initial attitude toward sharing knowledge through the Web site of the VLC is shown to be significantly associated with their intentions of performing such behavior. This is consistent with Bhattacherjee and Premkumar (2004)’s findings that IT users’ attitude toward IT usage will affect their intentions of subsequent IT usage. The positive and significant relationship between subjective norm and intention suggests that the opinions from teachers, friends or classmates help to form an atmosphere which facilitates the earnestness of joining in online discussion to acquire or apply knowledge. This result corroborates the subjective norm-intention relationship that was found by prior TPB studies conducted in various contexts. For example, a survey conducted by Bock et al. (2005) indicated that based on a sample of thirty organizations the subjective norm showed a significant influence on knowledge sharing intention.

The result also shows that Web-specific self-efficacy positively and significantly influences the intention. According to social cognitive theory (Bandura, 1997), an individual would strive for goals that he judges himself capable of attaining and avoid goals that he perceives as being too difficult. Self-efficacy theory also suggests a positive link between users’ a priori perceived level of self-efficacy and their desire to engage in a particular activity (Hunton & Beeler, 1997). The current finding echoes with Hunton and Beeler’s (1997) finding which suggests a significant and positive association between user self-efficacy and the desire to participate in target activities. Hence, an individual with beliefs in his efficacy will set more difficult goals for himself and thus mount the level of motivation needed to reach higher levels of performance. Other positive and significant relationships (hypotheses 5 and 7) observed in this study suggest that Web-specific self-efficacy and knowledge creation self-efficacy are important antecedents leading to knowledge sharing behavior. This finding is consistent with the findings of Marakas et al. (1998) and Compeau and Higgins (1999) which show that task-specific self-efficacy exerts strong influence on actual behavior of the target task.

Social network ties appear to be positively and significantly associated with knowledge sharing intention. This implies that once an individual build up relationships with other VLC members that he feels comfortable to share his ideas, thoughts, or story, his intention of performing such behavior will be stronger. This result supports previous findings of VC (Chen, 2007) and VLC (Yang et al., 2007) studies which suggest that social interaction ties help to enhance community members’ intention to share knowledge online. Finally, a significant relationship is observed between knowledge sharing intention and the actual behavior. This is consistent with the result of a TPB-based electronic commerce study (Pavlou & Fygenson, 2006) which finds a significant association between online get-information intention and get-information behavior. The current results prove the explanatory power of TPB in the VLC context.

Against prior expectation, the relationship between knowledge creation self-efficacy and knowledge sharing intention does not prove previous conjecture. People engage in volitional behavior because it is self-relevant; it has subjective meaning and importance to the individual (Rejeski et al., 2005). Contributing knowledge online can be a semi-volitional behavior in VLCs given that performing online knowledge sharing weighs heavily in performance assessment. Therefore, a possible explanation of this result is that learners consider sharing knowledge online as a mean to achieve the end of earning the course credit, those with low self-efficacy in expressing ideas, applying knowledge, or responding messages may still have high intentions to strive for the goals.
Conclusion and Implications

Enhancing learning performance through the setting of virtual learning communities has become an emergent trend in educational institution as well as business organizations. Although the target subject in this study includes senior college students and MBA students who took courses conducted in the VLC setting, the research findings can be applicable to knowledge workers in business organizations and communities of practice. For business organizations, the goals of putting into practice such training policy are to train their personnel in a more efficient manner. For educational institutions, cultivating learners’ capabilities in exploring and exploiting knowledge from the unlimited resource in a virtual learning community can help them not only learn more efficiently, but also learn ‘how to learn’ in the digital economy era. However, business managers or instructors need to foster an atmosphere of community where individuals hold positive attitude about interacting with strangers on Internet. Moreover, prerequisite training for using the VLC website functions and knowledge creation capabilities in the VLC setting is also required to achieve its intended goals.

This study suffers from a major limitation. We focus on learners’ capabilities, attitude, centrality in the social network, and opinions from others in the network to explain the social behavior. Other virtual community studies (e.g., Chen, 2007) have found that website quality is an important variable influencing users’ intentions to participate in VC activities. Prior IS research suggested that the website quality encompasses three dimensions-information/knowledge quality, system quality and service quality. Therefore, further studies are suggested to explore the impact of the VLC website quality, including knowledge quality and system quality, on learners’ knowledge sharing behavior.

Although some VLC studies have examined the impact of learners’ attitude or self-efficacy alone on learning performance, to our best knowledge, little VLC research has provided a comprehensive theoretical model grounded on related theories. This research is one of the earliest VLC studies which consider the characteristics of virtual learning communities and integrate the concept of social network ties, theory of planned behavior and two types of self-efficacy to advance the understanding of online knowledge sharing behavior in the VLC setting.

Acknowledgements

We would like to express special thanks to the two anonymous reviewers for their thoughtful comments and constructive suggestions. This work is supported by National Science Council, Taiwan under grant NSC 96-2520-S-231-001.

References


### Appendix A. Measurement items

<table>
<thead>
<tr>
<th>Construct</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attitude Toward Knowledge Sharing</strong></td>
<td></td>
</tr>
<tr>
<td>ATT1</td>
<td>I believe that sharing knowledge with others is useful for learning new knowledge.</td>
</tr>
<tr>
<td>ATT2</td>
<td>I believe that sharing knowledge with others will help me learning more efficiently.</td>
</tr>
<tr>
<td>ATT3</td>
<td>I believe that sharing knowledge with others gives me the opportunity to learn from more people.</td>
</tr>
<tr>
<td>ATT4</td>
<td>I believe that sharing knowledge with others is useful for enhancing my learning performance.</td>
</tr>
<tr>
<td><strong>Subjective Norm</strong></td>
<td></td>
</tr>
<tr>
<td>SN1</td>
<td>My teacher thinks that I should share my knowledge with other members in the virtual learning community.</td>
</tr>
<tr>
<td>SN2</td>
<td>My friends think that I should share my knowledge with other members in the virtual learning community.</td>
</tr>
<tr>
<td>SN3</td>
<td>My classmates think that I should share my knowledge with other members in the virtual learning community.</td>
</tr>
<tr>
<td><strong>Knowledge Creation Self-Efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>KCSE1</td>
<td>I feel confident clearly expressing my ideas in the virtual learning community.</td>
</tr>
<tr>
<td>KCSE2</td>
<td>I feel confident responding others’ messages in the virtual learning community.</td>
</tr>
<tr>
<td>KCSE3</td>
<td>I feel confident articulating my ideas into written, verbal or symbolic forms.</td>
</tr>
<tr>
<td>KCSE4</td>
<td>I feel confident applying my knowledge to help others resolve their problems.</td>
</tr>
<tr>
<td><strong>Web-Specific Self-Efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>WBSE1</td>
<td>I feel confident acquiring sufficient resource and knowledge from the Web site of the virtual learning community.</td>
</tr>
<tr>
<td>WBSE2</td>
<td>I feel confident efficiently utilizing the resource provided by the virtual learning community, such as discussion forum, recommended articles and e-papers.</td>
</tr>
<tr>
<td>WBSE3</td>
<td>I feel confident posting messages in the Web bulletin board.</td>
</tr>
<tr>
<td>WBSE4</td>
<td>I feel confident browsing information over the Web site of the virtual learning community.</td>
</tr>
<tr>
<td><strong>Social Network Ties</strong></td>
<td></td>
</tr>
<tr>
<td>SNT1</td>
<td>I maintain close social relationships with some members in the virtual learning community.</td>
</tr>
<tr>
<td>SNT2</td>
<td>I intensively exchange ideas with some members in the virtual learning community.</td>
</tr>
<tr>
<td>SNT3</td>
<td>I know some members in the virtual learning community on a personal level.</td>
</tr>
<tr>
<td>SNT4</td>
<td>I have frequent communication with some members in the virtual learning community.</td>
</tr>
<tr>
<td><strong>Knowledge Sharing Intention</strong></td>
<td></td>
</tr>
<tr>
<td>INT1</td>
<td>I intend to frequently share my ideas with other members in the virtual learning community.</td>
</tr>
<tr>
<td>INT2</td>
<td>I will always provide my know-where or know-whom at the request of other members in the virtual learning community.</td>
</tr>
<tr>
<td>INT3</td>
<td>I will try to share my expertise from my education or training with other members in an effective way.</td>
</tr>
<tr>
<td>INT4</td>
<td>I intend to share my resolution of others’ problems to earn more course credit.</td>
</tr>
<tr>
<td><strong>Knowledge Sharing Behavior</strong></td>
<td></td>
</tr>
<tr>
<td>BEH1</td>
<td>I usually spend a lot of time sharing knowledge with other members in the virtual learning community.</td>
</tr>
<tr>
<td>BEH2</td>
<td>I usually actively share my knowledge with other members in the virtual learning community.</td>
</tr>
<tr>
<td>BEH3</td>
<td>I usually involve myself in discussions of various topics rather than specific topics.</td>
</tr>
<tr>
<td>BEH4</td>
<td>I usually respond to others’ comments on my messages.</td>
</tr>
</tbody>
</table>
Applications of a Time Sequence Mechanism In the simulation Cases of a web-based Medical Problem-Based Learning System

Lih-Shyang Chen¹, Yuh-Ming Cheng², Sheng-Feng Weng¹, Yong-Guo Chen¹ and Chyi-Her Lin³

¹Department of Electrical Engineering, National Cheng Kung University, Taiwan // TEL: +886 6 2757575
²Department of Computer Science and Information Engineering, Shu Te University, Taiwan // TEL: +886 7 6158000
³Dean of Medical School, National Cheng Kung University hospital, Taiwan // TEL: +886 6 2353535
chens@mail.ncku.edu.tw // cymer@mail.stu.edu.tw // afeng@mirac.ee.ncu.edu.tw // tyk@thmu.edu.tw // neonate@mail.ncku.edu.tw

ABSTRACT
The prevalence of Internet applications nowadays has led many medical schools and centers to incorporate computerized Problem-Based Learning (PBL) methods into their training curricula. However, many of these PBL systems do not truly reflect the situations which practitioners may actually encounter in a real medical environment, and hence their effectiveness as learning tools is somewhat limited. Therefore, the present study analyzes the organization of a computerized PBL teaching case, and considers how a clinical teaching case can best be presented to the user. Specifically, this study attempts to develop a web-based PBL system which emulates the real clinical situation, in which a practitioner will ascertain the true nature of a patient’s condition over a period of time, by introducing the concept of a “time sequence” within each teaching case. The proposed system has been installed in the medical center of National Cheng Kung University in Taiwan for testing purposes. The feedback received confirms the effectiveness of the training system in allowing users to apply their basic medical knowledge to real clinical cases, and in identifying their areas of misconception in the course of a medical diagnosis.

Keywords
E-Learning in medical education, Problem-Based Learning, Computer education, Educational technology

Introduction

Recently, many medical schools and centers had adopted a variety of approaches in an attempt to address various problems of the traditional medical training methods. In particular, as the Internet becomes prevailing, one such approach has been to introduce computerized Problem-Based Learning (PBL) techniques using a browser (DxR, 2006; Norman, & Schmidt et al., 1992). (It is acknowledged that the word “PBL” may be regarded by some as controversial. Nevertheless, its use is retained by the current authors since it is not the main concept addressed within the present paper.) In the medical field, this approach involves the presentation of clinical cases as a means of learning basic medical and clinical science (Bryce, & King et al., 1999; Joines, & Barton, 2000). Previous studies have confirmed that implementing PBL techniques on computer networks provides an effective training tool for physicians or students in medical schools, and enables these users to acquire the necessary knowledge and experience required to perform accurate diagnoses in their actual clinical practices (Diana, & Dolmans et al., 2002; Barrows, & Tamblyn, 1980). The basic three steps (Barrows, 1985) of a computerized PBL system within the medical training field are as follows: (1) Students are puzzled and challenged by the patient’s problems that are given all at once. (2) The system will provide some selection types of questions for the students to get more information needed, such as questioning the patients and ordering some laboratory tests. After the students’ selection, the corresponding answers to the questions will be provided so that the students can get more information about the patient’s status. For instance, the system enables the users to select any kind of laboratory tests and provides the answers to them to emulate the real clinical situation. (3) The students can repeatedly interact with the system interactively and work on their ways in an attempt to diagnose and cure the patient’s problem. Current web-based PBL systems do help students’ learning to some extents, they suffer from several serious drawbacks, which degrade their effectiveness as learning aids. For example, they do not truly reflect the typical clinical situation, in which a patient’s condition may evolve over time. This drawback basically arises from the manner in which the data related to a particular teaching case is presented to the user.

Consequently, this present study considers how a computerized PBL teaching case should best be presented to the user such that it serves as a truly effective learning tool (Gülseçen, & Kubat, 2006; Cheng, & Chen et al., 2003;
Segers, & Bossche et al., 2003). The theoretical aspect of the computerized PBL has been studied extensively in (Herrington, & Oliver, 2006; Herrington, & Oliver et al., 2003) which described it as online authentic learning. This study focuses on how the theoretical concepts described above can be designed and implemented in a real system for a specific learning environment of the medical education. The study also discusses the implementation of a prototype system, which is currently being used to test the approach proposed. In Section 2, we briefly describes the concept of the teaching case template adopted by the proposed system. Furthermore, the background to the most commonly available computer-based or web-based PBL systems (Rolf, & Uno, 2003; Kimball, & Shin et al, 2003; Hjelm, & Lee et al., 2002) is discussed, and their relative drawbacks are analyzed. The adoption of a time sequence concept within the computerized PBL teaching cases as a means of addressing these limitations is also discussed. Section 3 adopts the concept of a time sequence in proposing various storage and presentation formats for medical teaching cases. In order to really demonstrate the idea of the time sequence, we also implement the idea in a web-based computerized PBL teaching case system that allows users to simply use a WWW browser to browse the contents of the system. Section 4 discusses the architecture of the proposed system and its implementation, and Section 5 describes the results of our experiments with the proposed system. Finally, we give some brief conclusions of the present study in Section 6.

![Figure 1. The layout of the main HINTS windows](image)

Relevant background

Description of HNTS

For explanation purposes, we first briefly give an example of how HINTS works. The main window of HINTS is shown in figure 1. The left panel of the window is the display of the teaching cases organized as a tree-structure. The right panel of the window is the display of the sections of a given teaching case. The student is supposed to select a case first, then select one of the sections for browsing the selected case. The contents of the corresponding section will be displayed at the center portion of the window. The HINTS is basically a PBL system in which the student is challenged by the initial status of a patient and works his way toward a correct final diagnosis (by browsing through several sections) followed by patient management until the patient is completely cured. In other words, HINTS is a simulation system that allows the users to diagnose and cure the virtual patient in the system. In the HINTS, the student is first challenged by the patient’s chief complaints and reads the patient’s basic information section where the patient’s height, weight, age and so on are described. Then, in the present illness section the system will present to the student many questions to ask the patient. The student is supposed to select which questions are critical to ask in order to reach a correct diagnosis for this particular case and the corresponding answer to the selected question will be shown by the system. In the Physical Examination (PE) section, the system presents a patient’s figure as
shown in figure 2(a). The student is supposed to first select one of the PE types: Inspection, Palpation, and Auscultation followed by clicking on the portion of the human body where the selected PE should be carried out. The system will respond with the associated multimedia information, such as the heart beat sound for the auscultation of a certain point on the patient’s chest, and the picture of the skin for the inspection of a certain portion of the patient’s body. In the laboratory section, the student is supposed to enter a hypothesis – a disease name followed by clicking on various test items as shown in figure 2(b) to test against the hypothesis. The system will display the associated test results, such as Computed Tomography (CT) images or Ultra-sound images. The same process can be repeated over and over again. In the course of this diagnosis procedure, the student can go back and forth among these sections until he or she gets enough information and feels comfortable to give a correct final diagnosis followed by the patient management section.

2(a) the section on Physical examination

2(b) the section on Laboratory tests

Figure 2. The presentation of the Physical examination and Laboratory test
The HINTS is a complicated multimedia simulation system that allows students to deal with the virtual patient in the system to emulate dealing with a real clinical case.

The concept of teaching case templates

In order to explain how the developed prototype system is implemented, it is first necessary to introduce the concept of the PBL teaching case template. The developed PBL teaching case system is essentially a multimedia CAI (Computer Aided Instruction) system. Generally speaking, such systems comprise three basic models, namely a knowledge model, a student model, and a tutor model (Clancey, 1987).

Knowledge model

The knowledge model is a database containing the knowledge of specific topics which an expert in that particular domain could reasonably be expected to possess. The data stored within the knowledge model represent the knowledge which is to be taught to the student, and should ideally be stored in an abstract fashion such that the model is capable of dealing with different learning situations in a flexible and intelligent manner. Therefore, the current knowledge model utilizes a number of “teaching case templates” (Cheng, & Chen et al., 2003) to form an abstract model of the domain knowledge. In other words, the templates outline the main contents of the various teaching cases, and serve as the directories of each case. A typical medical teaching case might well include the following sections: (1) basic personal information such as the patient’s age and gender, (2) a brief case history, (3) reported complaints, (4) results of physical examinations, (5) findings, (6) the diagnosis, (7) relevant cases, (8) discussions, (9) comments, and (10) learning points (i.e., learning goals of the teaching case). The details of all sections of the template can be presented via hypermedia techniques. The authors of the PBL system are provided with the tools required to create a case template or select a case template from an existing case template database, in which each template has a defined name and an appropriate set of titled sections, e.g. History, Diagnosis, etc as described above. If necessary, each of these sections can be further partitioned into several sub-sections. It is noted that a single teaching case template may be used by many teaching cases of the same type. Each teaching case within the system identifies the particular case template which it uses.

Student Model

For each student this model includes the following: (1) student profile, i.e. basic information about each student, including name, ID, department, student level, etc, (2) records of the case categories which that student is interested in, (3) records of the cases which that student has reviewed, and (4) performance evaluation results.

Tutor Model

This model mimics a human tutor, and is responsible for managing the overall learning environment (Cheng, & Chen et al., 2007).

Discussion of the drawbacks of the current computerized PBL teaching case systems, and an introduction to the concept of time sequence in a teaching case

In many computer-based or web-based PBL teaching case systems (Rolf & Uno, 2003; Dxr, 2006; Giardina, & Oubenaissa et al., 2002) as briefly described in three steps before, the user is first presented with basic patient information and details of the patient’s principal complaints. He is then provided with the means to specify the sources from which he wishes to get relevant information to the case, e.g. history taking information, laboratory examination items, etc. The detailed information, as the answer to the selected items, is then retrieved and presented to the user, who can then select whichever information he feels to be necessary to make an appropriate final diagnosis. Once the user has provided his diagnosis, the system will assess his performance in completing the teaching case.
In this type of computerized PBL teaching case, all of the relevant information relating to a particular case is lumped together within the system. However, in practice, it is far more likely that this information will be accumulated gradually over a long period of time. In other words, conventional computerized PBL techniques do not include the concept of a “time sequence” of events within their teaching cases. For reasons of convenience, this type of teaching case will be referred to as a “lump-teaching case” throughout the remainder of this paper. Although this type of teaching case may be of some value for training purposes, it is nevertheless rather unrealistic, and suffers from the following drawbacks:

- In many cases, a medical practitioner is unable to make a final diagnosis after the patient's first visit. Several patient visits may be needed before the final diagnosis can be made. During the on-going course of diagnosis and treatment, the results of the physical examinations, and the patient’s response to the prescribed medicine, may lead the practitioner to consider or to discount various diagnoses. Each stage in the medical decision-making process has its own reasoning and inference procedures. To accurately represent the medical situation described above, and to emulate the associated decision-making processes, it is important that the PBL teaching case presents the various events in an appropriate time sequence. The overall diagnosis and treatment of a patient may well involve a series of treatments and physical examinations, and hence the diagnosis and treatment process can be regarded as a series of events taking place at different points along a time axis instead of happening all at once as described in most of the current computerized teaching cases found on the web.

- There are some other situations where the concept of time sequence in a teaching case is necessary. For instance, in many actual cases, local regulations may prevent the medical practitioner from scheduling physical examinations at will, or at least not unless the patients are willing to meet the cost of these examinations from their own pockets. For example, in Taiwan, the National Health Insurance Scheme, which is administered by the central government, does not permit a medical practitioner to schedule certain physical examinations until other preliminary tests have first been performed to confirm that the further examination is truly necessary. This policy exists to avoid wasting limited insurance fund resources, and furthermore, to protect the patient’s general health, since some examinations may cause physical discomfort, or even physical harm. Although the example cited above refers specifically to Taiwan, the same situation is common in many other countries which operate similar health insurance policies, or to cases where private insurance companies will meet the cost of the examination.

- From a purely clinical perspective, the physical examinations should be sequenced in such a way that the cheaper, non-invasive, and safer examinations are performed first. The more expensive, invasive and riskier examinations are performed at a later stage in the course of treatment if and when they are called for. In most computerized teaching cases which lump all of the information together, it is difficult to determine whether or not a student is following the correct line of thought in scheduling the appropriate examinations in an appropriate sequence.

The main point of the preceding discussions is that each of the events within a course of treatment occurs at a discrete time, and hence this sequence of events and the corresponding accumulation of the thinking processes should somehow be emulated within the computerized PBL system. The theoretical aspect of the computerized PBL has been studied extensively in (Herrington, & Oliver, 2006; Herrington, & Oliver et al., 2003) which described a practical framework for the design of learning environments. The framework includes nine situated learning design elements. The present study tries to incorporate the nine design elements of that framework through the concept of the time sequence in the proposed teaching case system. The system provides “time-sequence teaching cases” in which each individual event is recorded, and presented, as part of an overall sequence of events distributed along a time axis.

**Different formats of PBL teaching cases**

As indicated above, the clinical teaching case materials can either be arranged in a lump teaching case format or in a time-sequence teaching case format. From a more abstract point of view, for each format just described, there are two further formatting issues to be considered, namely the format in which the teaching materials are stored in the computer system, and the format in which these materials are then presented to the end user. For the sake of convenience, these formats are referred to as “the storage format” and “the presentation format”, respectively. The relationships of the storage and presentation formats are depicted in figure 3. This two-layer arrangement of the teaching materials allows for the de-coupling of the storage and presentation formats, e.g. a teaching case can be stored in one format and presented to the end user in the same format, or it can be re-formatted using web-based or
artificial intelligence technologies at run time and presented to the user in a manner which best meets his training needs. Thus, this arrangement greatly enhances the flexibility of the system and maximizes the usage of the data stored within it.

In the same way that there exists two storage formats, i.e. lump-storage and time-sequence, there also exists two presentation formats, i.e. lump-presentation and time-sequence presentation. Combining these various formats results in a total of four different formats, namely lump-storage lump-presentation, lump-storage time-sequence-presentation, time-sequence-storage lump-presentation, and time-sequence-storage time-sequence-presentation. The appropriate storage format for a particular case is specified by the author at the time of editing the teaching case materials into the system. However, the presentation format, which dictates how the materials within the teaching case will be displayed to the user, can be specified at run time. The various storage-presentation format combinations are described in more detail in the following paragraphs.

As described in section 2, the teaching case template represents the meta data of a teaching case, i.e. it is the basic model of the teaching case. The case template is also used to model the information associated with each patient visit, and is referred to as the “visit page”. A typical visit page includes several sections, e.g. history taking, chief complaints, physical examinations, tentative diagnoses, final diagnosis, treatment, etc. It is noted that the notation of “patient visits” is adopted as an abstract means of reflecting the time sequence of the various events which may take place during a short period of time for a particular patient. In other words, the term “second visit” may actually represent the second day of a patient’s stay in the hospital or a second separate visit to the medical practitioner.

In the lump-storage lump-presentation format, all of the case materials are stored within a single visit page. The user can then order whichever examinations he feels necessary to make a diagnosis, and can view the examination results immediately. The user can repeat this procedure continuously until he feels confident about making a final diagnosis. This format is similar to that provided by the commercially available computer-based or web-based teaching case systems described previously, and is particularly suitable for novice users who are developing their diagnosis skills. This format is also suitable for the majority of medical image teaching cases (Wong, & Hoo, 1999; Dev, 1999) in which the diagnosis procedure can be performed in “one shot”. It is noted that the adoption of this particular format greatly simplifies the author’s task in composing the teaching case.

In the time-sequence-presentation format, the user specifies whichever examinations he feels to be necessary and submits some tentative diagnoses. However, the results of the examinations are not made available to the user until he has completed the tentative diagnoses section. Once he has done so, the system automatically displays the next patient visit page, and the user can view the examination results via the “the results of the examinations” section, which is the first section in the case template for this visit page. This procedure emulates the actual situation encountered in a genuine clinical setting, and can be repeated by the user until he feels able to provide a final diagnosis. During this iterative process, the system records the user’s actions at each step such that his thinking
processes can be analyzed. For example, the system records the number of patient visits required before the user provides a correct diagnosis. In the lump-presentation format, the user is able to order several examinations at the same time, but there is no means of verifying whether he intends to perform all of them at one time (i.e., in one visit), or if he is planning to schedule them over several patient visits. Clearly, these two scenarios imply different thinking processes. In the time-sequence-presentation format, the examination orders are set out clearly along a time axis, and consequently it is far easier to track and analyze the user’s thinking process.

In the lump-storage time-sequence-presentation format, the user can order specific examinations and check the results at the next visit in an iterative fashion until he is in the position to make a final diagnosis. In this arrangement, although all of the necessary data are stored in the system in a lump format, the system still allows the user to interact with the system over a series of patient visits. This format is particularly useful when the correct final diagnosis of a particular teaching case can be reached by following a variety of different thinking processes. In other words, this format tests the ability of the user to order the correct examinations and to provide the correct diagnosis, rather than his ability to simply follow a single “correct” path of reasoning. This format is also suitable for the situation where all of the data in a teaching case can be presented in “one shot”.

In the time-sequence-time-sequence-presentation format, a typical teaching case consists of several visit pages of information, where each page represents the data relating to a particular patient visit at a certain point in time. This format most closely resembles the genuine clinical situation, such as the patient’s condition keeps changing over time, and gives the student a strong feeling for the actual clinical environment. Furthermore, this format is particularly suitable for those medical cases which have a standard (or predetermined) diagnosis procedure and corresponding course of treatment.

In the event that a teaching case with this particular format is beyond the ability of a novice, the time-sequence-storage lump-presentation format can be used to enable the user to browse the case in a lump-presentation format. In this learning environment, the original teaching case in a time-sequence-storage format is simply pre-processed into a corresponding lump-storage format by removing the time factor. The teaching materials are then compiled and presented to the user in a lump-presentation format.

**Implementation issues**

The implemented system contains editing tools which allow the authors to edit both the case templates and the teaching case materials. We use teaching case template as the schema of teaching case database. The case templates and the teaching case materials are then stored in the teaching case databases. When a particular case is activated, the relevant data, i.e. the associated case template and the applicable case contents, are retrieved from the databases and made available for presentation processing in accordance with the required presentation format.

As described previously, the proposed system allows for four different types of data storage and presentation formats. Therefore, it is necessary to cater for various situations at run time. One way to solve the presentation problem is to separate the mechanism which provides a specific effect for presentation purposes from the policy which determines when that specific effect should take place. In other words, the system should incorporate a series of mechanisms which can provide all the possible presentation effects which might be required without actually having to decide which particular effect should be activated. The policy which decides which presentation mechanism should be activated is established through a completely separate decision-making process. Once the mechanisms themselves have been defined, this decision-making policy may be specified at any time after the case has been installed within the system, or even at run time.

In the implementation of the proposed system, a rule table is used to specify the policy as to how a case should be presented, while switch flags are employed as mechanisms to present the same set of teaching materials in different formats or manners at run time. In the implementation, the switch flags are in fact several bits each of which controls a certain presentation function to be turned on or off so that the specific effects can take place. This is because the implementation of a presentation format involves many different switch mechanisms in the presentation flow of a teaching case. Each switch mechanism can be controlled by a switch flag. In other words, a presentation format corresponds to several switch flags each of which controls the presentation mechanism at run-time.
Regarding the teaching case presentation policy, the implemented system has two databases:

- A user profile database which records each individual user’s major (e.g. Internal Medicine) and his level of ability (e.g. “Intern”, “Specialist”, “Medical School Student”, etc.).
- A view-rule database comprising the following fields: (1) user level, (2) user major, (3) case template ID, (4) section name, (5) presentation format (i.e. lump-presentation or time-sequence-presentation), and (6) a switch flag which records how each section of the teaching case should be presented. Items 1 to 5 are used as indexes to retrieve the appropriate switch flag, which then specifies how the corresponding section is to be presented.

The user profile database must be populated by the system administrator, and the view-rule database by the authors, before the teaching cases can be published on the website. The overall architecture of the proposed system is shown in Figure 4. It is noted that all of the multimedia documents associated with the various teaching cases are stored in a teaching case database within the system in order to simplify their management.

When a user accesses the web browser and logs into the system, the system will first retrieve that individual’s major and default user level from the user profile database. When the user activates a teaching case for interactive browsing, the system will use the user’s level and major, the case template ID (stored in the case), and the specified presentation format, and based upon this information will then determine the appropriate switch flags from the view-rule database such that the data within each section is presented in the appropriate format. As shown in Figure 5(a), the system presents all of the section names associated with the particular case template in the form of push buttons on the right side of the screen. Meanwhile, in a sub-window on the left side of the screen, all of the teaching cases which the user can access and browse are displayed in the form of a hierarchical structure. It is noted that this sub-window can be closed by the user at any time if it becomes necessary to free up additional screen space to display the contents of the current teaching case (as shown in Figure 5(b)). When a particular section is activated, the system acts in the same way as a web service application system responding to a web browser’s request, and retrieves the appropriate contents from the teaching case database, references the corresponding switch flag, and then presents the contents of that particular section in the required format.

**Figure 4. System Architecture**

Figures 5(b) and 5(c) show typical screens relating to the second and third visits. From the discussions above, it can be seen that the switch flag corresponding to each section governs how that particular section should be presented at run time. As a result, a single teaching case can either be presented in a lump-presentation format or in a time-sequence-presentation format. Furthermore, for a given teaching case, this mechanism enables the system to present a different subset of the same teaching materials to different users. For example, if the user is a medical school
student, when he accesses the tentative diagnosis section of a case, the system can show six potential tentative diagnoses and can then prompt the user to select one of them. However, for an intern user, who is reasonably expected to possess a greater medical knowledge and more clinical experience, the system could prompt the user to enter his own tentative diagnosis -- spell out the tentative diagnoses, i.e., a fill-blank type question as opposed to a selection type question. In other words, for a single section in a teaching case template, the system has the ability to present the relevant teaching materials in many different ways. This is a highly valuable by-product of the system mechanism which was originally designed solely to implement the time sequence concept within teaching cases.

(a) System Screen Layout

(b) Second Visit
Experimental results

HINTS has been up and running in the medical center of National Cheng Kung University for 3 years. Currently, they have many practical training courses in the medicine department. The objective of the courses is to learn the skills of how to face a real patient, order laboratory test, make a diagnosis, and give treatment to the patient. We use these teaching cases in the HINTS to emulate the real cases. Currently, there are 50 teaching cases in the system for the clinical practice training of the medical school students. The system was implemented on Microsoft Internet Information Server version 6.0 (IIS) using both Microsoft ASP (Active Server Page).Net, and SQL database technologies. Students can simply use the WWW (World Wide Web) browser to login to the HINTS and browse the teaching cases at the computer center in the medical school or at home, if a wide bandwidth Internet connection is available. Before the mechanism of time sequence was installed in the system, there were more than 200 medical school students from the 5th to 7th grades (equivalent to 1st to 3rd grade students in a medical school in the American medical education system) had experienced in using the system.

1. Understand all items
2. Define the problem
3. Analyses the problem
4. Synthesize
5. Define learning objective
6. Self-study
7. Report back

1-5: pre-discussion  7: post-discussion

Since the main purpose of this study is to investigate how a teaching case that have multiple visits should be implemented in a real computerized simulation learning environment, such as the HINTS, we need to find out whether the designed time sequence in teaching case is really effective in particular for the training of the thinking
process. The participants of the experiments are fifty medical students ranging from Grade 5 to Grade 7. The learning process is structured according to the seven-jump procedure (Segers, & Bossche, 2003), as shown in figure 6.

Procedures

Prior to the experiment, we demonstrated the HINTS operation including the time-sequence-presentation to all the students and made sure the students could use HINTS without difficulty. At the beginning (the first classroom lecture for a case), the instructor gave the students some background knowledge about the particular case in the classroom. There were 6 teaching cases used in the experiments: 2 simple ones, 2 mediocre ones, and 2 difficult ones. The students browsed through these cases sequentially. Then, the students browsed through the case including reading through the basic information and chief complaints sections in the HINTS, specifying which part of the patient body should be examined, what questions should be asked, and what laboratory test should be ordered to get more information and insight about the patient’s status from the HINTS. They went through several patient visits as time went on, and finally gave their final diagnoses for the exercise.

Furthermore, right after the students finished their browsing of the six teaching cases, they were asked to fill in an on-line questionnaire to collect the system evaluation data. The questions in the questionnaire were designed simply to see whether our design and implementation of the time sequence mechanism in teaching case was effective or not. In order to obtain a reliable survey result, expert validity was adopted in this study. The questions in the questionnaire were reviewed, revised, and eventually approved by 6 experts including three medical education experts, one statistics expert, and two information science experts. In the second classroom lecture, the instructor and students discussed their results computed by the HINTS and experiences”.

Results

Table 1 uses a five-point Likert Scale and summarizes the student responses to a series of questions posed by the current researchers. The students have also made a variety of comments after using the various teaching cases in the system. The principal results of this study may be summarized as follows (some of the results are provided by the instructors).

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage of Respondents % (n=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice can improve the individual learning process?</td>
<td>Strongly agree (5) 28</td>
</tr>
<tr>
<td>The use of multiple visits can simulate the real clinical situation?</td>
<td>6</td>
</tr>
<tr>
<td>The operation of the multiple visits cases was easy?</td>
<td>6</td>
</tr>
<tr>
<td>The multiple patient visit cases enable you to learn and trace more about the patients’ condition?</td>
<td>12</td>
</tr>
<tr>
<td>The system can improve the capability of the clinical service?</td>
<td>12</td>
</tr>
<tr>
<td>The limited number of patient visits makes the case more interesting and more challenging?</td>
<td>28</td>
</tr>
<tr>
<td>Do you think the time-sequence-presentation case better resemble the actual clinical setting and provide better feeling for what actually takes place in an authentic clinical environment than the lump-presentation case?</td>
<td>36</td>
</tr>
</tbody>
</table>
• All of the students feel that the teaching cases presented using the time-sequence-presentation format approach better resemble the actual clinical setting, and provide a better feeling for what actually takes place in an authentic clinical environment than the cases presented using the lump-presentation format.

• All of the students believe that the system can significantly develop their skills for dealing with real clinical cases. Furthermore, when a patient’s condition evolves over time, it is felt that the concept of a time sequence within the teaching case is particularly meaningful especially, for those complex teaching cases.

• All of the case templates within the implemented system include a learning points section, which can be accessed by the user after he has completed a particular teaching case. All of the students believe that this is a highly useful function.

• 72% of the students think that if the system limits the number of patient visits a student can have during the diagnosis procedure of a teaching case, it will cause the teaching case to become more interesting and more challenging.

• Through the analysis of the logs of the HINTS server, the students spent 58 minutes for browsing a case with the time-sequence-presentation while they spent 29 minutes for a case without the time-sequence-presentation on average. When we interviewed the students for collecting their opinions about the system, they expressed that it took a long time to finish a case. They thought that the reasonable time duration for browsing such a case should be around 45 minutes. This suggestion is useful for the future development of the teaching case with the time sequence mechanism built-in.

• The time-sequence-presentation of a teaching case enables the teacher to check whether the student has ordered the appropriate physical examinations at the appropriate stage of the diagnosis process, i.e. from cheap, non-invasive, and safe examinations during the initial stages, towards more expensive, invasive, and riskier examinations towards the latter stages. Additionally, the system also presents statistics relating to the provision of correct and incorrect answers, and therefore allows the teacher to analyze a student’s progress in adhering to an appropriate diagnosis procedure as his training program continues.

Conclusion

This study has incorporated the theoretical framework of online authentic learning (Herrington, & Oliver, 2006; Herrington, & Oliver et al., 2003) and has proposed the concept of a time-sequence in the storage and presentation formats of computerized PBL teaching cases. An interactive PBL teaching case system has been implemented which introduces this concept by separating the mechanism which provides the required presentation effect from the policy which specifies which mechanism to activate. Experimental results have indicated that users feel strongly about the way in which cases are presented to them.

Furthermore, the users have confirmed that the teaching cases are more interesting and challenging when the time sequence concept is used, and that these cases provide a better feeling for the actual clinical environment which they will work within in the future. Using the same framework which enables the system to provide different presentation formats, the system can also be customized to meet an individual user’s particular training requirements. Although this paper has presented a series of preliminary experiments involving the time sequence concept, it is acknowledged that the long term impact of the proposed system for medical education as a whole still remains to be seen.

Acknowledgement

The authors wish to thank the medical students of National Cheng Kung University for participating in the evaluation process and feedback of HINTS. This project was supported by a grant from National Science Council, Executive Yuan ROC (NSC 92-2524-S-366-001).

References


The information commitments toward web information among medical students in Taiwan

Jyh-Chong Liang
Department of Holistic Wellness, Chin Min Institute of Technology, Taiwan // aljc@ms.chinmin.edu.tw

Chin-Chung Tsai
Graduate School of Technological and Vocational Education, National Taiwan University of Science and Technology, Taipei 106, Taiwan // cctsai@mail.ntust.edu.tw

ABSTRACT
The purpose of this study was to investigate medical students' information commitments toward online web medical information. The sample consisted of 534 volunteer medical university students with internet experience from 6 medical universities in Taiwan. This study used the Information Commitment Survey (ICS) for an investigation of medical students' standards of judging online information, and their search strategies on the Web. The results showed that the ICS was deemed to be sufficiently reliable for assessing medical students' information commitments toward Web-based information. It was also found that gender difference existed in the students' usage of certain searching strategies. Another finding was that students with more Internet experience tended to utilize the 'elaboration' searching strategy while being oriented towards employing quite 'mixed' standards for judging online information. The results also revealed that medical students held less sophisticated information commitments than university students in general, particularly when comparing to the female students. The findings strongly suggest that medical students need additional training for using the Internet to seek information and that medical educators implementing Internet-assisted instructional activities need to find some methods to help medical students, especially the female students, to develop more sophisticated information commitments.

Keywords
Information commitment, Medical studies, Gender, Internet

Introduction
Internet technology has changed not only our way of living but also the methods of learning in educational practice. Internet use has increased progressively over the last decade, and information searching has become one of the most popular and frequent online activities (Nemoto et al., 2007; Meneses et al., 2005). One of the major usages of the Internet for learning is to search for information on the Web for some academic goals. Information seeking is the process by which users find and select proper information resources in order to enhance their knowledge level (Segev et al., 2007). The Internet has also dramatically altered the worldwide use of academic publications for medical education as rich resources are available on the Internet for researchers and university students (Junnii, 2007). Tsai and Tsai (2003) have claimed that in the process of information searching, learners may use various types of searching strategies to obtain information they need on the Internet. Moreover, when judging the information they have found, they may employ various standards to evaluate its accuracy and usefulness (Wu & Tsai, 2005).

Nowadays health information is more and more readily available for users on the Internet with the rapid growth of medical information websites (Benigeri & Pluye, 2003). More recently, web-based resources supply regular professional updates, evidence-based patient care information, and medical problem solving (Vincent et al., 2006). Health and medical information is now one of the most frequently sought topics on the Internet (McMullan, 2006). One particular advantage of using the Internet for seeking healthcare information is the ease of acquiring information for a broad range of clinical subjects (Doyle, 2002). These web-based sites also provide medical students, patients and clinical physicians with a wide range of information relating to medical problem solved in general. Although there has been a sudden increase in the amount of medical or healthcare information available on the Internet, many studies have raised concerns about Internet sites that provide medical or healthcare information too soon after it has become available (Nemoto et al., 2007). Earlier studies have reported limited functions of the search engine’s usage by medical information seekers on the Internet (Suarez et al., 1997). The rapid growth of medical information on the Internet has resulted in questionable quality and potential dangers related to its flawed or unsuitable use (Benigeri & Pluye, 2003). Special attention needs to be focused on growing Internet sites with abundant medical information. For
example, Matthews et al. (2003) provided a scheme for testing Internet sites to identify those which provide scientifically accurate information while Sladek et al. (2006) tried to build up a search filter to select applicable information in the medical literature. In addition, most of the studies which assessed the quality, influence, usage or outcomes of the information on the Internet are related to websites in Western countries (Chu & Chan, 1998; Haddad & Macleod, 1999; Thomas & Kern, 2004; Rice, 2006; Voigt, 2007; Kari & Savolainen, 2007). With few researches exploring how precisely the users process the medical information on the Internet in an Eastern context, this study therefore aimed at investigating Taiwanese medical students’ online information judgments and searching strategies to address this research issue in an Eastern context.

Many universities now use the Internet, telecommunication technologies and multimedia programs as part of educational training in the classroom (Slotte et al., 2001). It is expected that medical students may frequently search/use web-based information for academic purposes. In medical schools, the Internet is now widely used by educators in instructional practice to help students acquire more knowledge and improve their learning outcomes (Slotte et al., 2001; Simmons et al., 2005). Appraising medical students’ usage and views of the Internet has also become an important research topic in medical education.

In medical schools, searching for medical information has now become a required component of clinical or research laboratories (Dikshit et al., 2005). While medical students are gradually taking advantage of the Internet in their learning, again, some situations based on our observations in Taiwan are especially worth gaining the attention of medical educators or teachers. For example, medical students may lack proper strategies to guide them to seek medical information through the Internet. Although they may feel overwhelmed by the amount of available online medical information that is relevant to their academic goals, most students might not fully comprehend this point, and they may depend on the Internet resources to provide accurate information without carefully ensuring the accuracy of the information they obtain. Our concern is for the lack of careful use of online information in medical decision-making by the students and even in the future when these students work in hospitals or medical-related settings. Few previous studies have addressed this important research issue.

To explore this issue, we adopted the concept of ‘information commitments’ for investigation. Information commitments are a set of evaluative standards in which Web users utilize in order to assess the accuracy and usefulness of Web-based materials, and these commitments are also relevant to searching strategies. Tsai (2004) proposed a theoretical framework for describing web users’ information commitments, including three aspects: standards for accuracy, standards for usefulness, and searching strategy. By interviewing 10 college students and two experts, two possible orientations for each aspect were revealed, yielding six possible factors of representing information commitments, including ‘multiple sources as accuracy,’ ‘authority as accuracy,’ ‘content as usefulness,’ ‘technical issues as usefulness,’ ‘elaboration as searching strategy’ and ‘match as searching strategy.’ Tsai (2004) concluded that the three information commitments, including ‘multiple sources,’ ‘content’ and ‘elaboration,’ were advanced information commitments while the others were considered as less sophisticated. An ‘Information Commitment Survey’ (ICS) consisting of these factors was developed by Wu and Tsai (2005), who found that learners’ evaluative standards of Web materials had a significant effect on their information searching strategies in web-based learning environments.

Wu and Tsai (2007) further assessed the information commitments of 1,220 college and graduate students in Taiwan using the ICS. The results showed that the ICS was sufficiently reliable for assessing students’ information commitments (Wu & Tsai, 2007). Wu and Tsai also found that male students were more oriented towards using the ‘match’ searching strategy to seek information on the web and that graduate students may be more oriented towards utilizing multiple sources to assess the correctness of the material in web-based learning environments than undergraduate students.

Although some previous studies have explored the growing trend of the utilization of the Internet to access healthcare information by ‘lay’ people (commons), health information seekers, adolescents, physicians and patients, consumers or college/master degree students (Chu & Chan, 1998; Nammacher & Schmitt, 1998; Dutta-Bergman, 2004; Gray et al., 2005; Escoffery et al., 2005; Bundorf et al., 2006; Schwartz et al., 2006; Lewis, 2006; Junni, 2007), the information commitments of medical students have not been carefully evaluated so far. In this study, the primary objective was to use the ICS to survey a group of medical students in Taiwan to probe their information commitments toward web-based information. To be more specific, the following questions were investigated in this study:
1. Is the ICS sufficiently reliable for assessing medical students’ information commitments toward web-based information?
2. What are the medical students’ information commitments?
3. Is there any gender difference in medical students’ information commitments?
4. What is the role of medical students’ Internet experiences in their information commitments?
5. What are the differences in the ICS scale scores between university students in general, reported by Wu and Tsai (2007), and medical school students in particular regarding information commitments?

Method

Sample

The participants of this study were 534 volunteer medical university students with internet experience (223 males and 311 females) from 6 medical universities in Taiwan. They constituted three major groups namely the medical group (department of medicine, department of dentistry and department of pharmacy, n=138), the nursing group (department of nursing and department of dental hygiene, n=209) and the health-care group (department of biological science and technology, department of health risk management, department of occupational safety and health, department of nutrition, department of occupational therapy and department of life science, n=187).

Instrument

This study used the Information Commitment Survey (ICS) developed by Wu and Tsai (2005) for an investigation of medical students’ standards of judging online information and their search strategies on the Web. Tsai (2004) proposed a theoretical framework for the ICS consisting of three aspects: (1) standards for accuracy, (2) standards for usefulness and (3) searching strategy, each of which included two scales. Therefore, the ICS consisted of six scales, that is, ‘multiple sources as accuracy,’ ‘authority as accuracy,’ ‘content as usefulness,’ ‘technical issues as usefulness,’ ‘elaboration and exploration as searching strategy’ and ‘match as searching strategy.’ Tsai (2004) also concluded that information commitments including ‘multiple sources,’ ‘content’ and ‘elaboration’ were more advanced information commitments while the others are considered less sophisticated. Wu and Tsai (2005, 2007) have found that the ICS is sufficiently reliable for assessing learners’ information commitments toward Web-based information. The items of the ICS were presented with bipolar strongly agree/strongly disagree statements on a six-point Likert scale (i.e., strongly agree, agree, somewhat agree, somewhat disagree, disagree and strongly disagree). Wu and Tsai in their paper suggested that the use of a six-point Likert scale should not only avoid neutral responses but also differentiate students’ variations of agreement in proper detail. A brief description of the six scales is presented below (For more details, please refer to Wu and Tsai, 2007):

1. **Multiple sources as correctness scale (Multiple sources):** measuring the extent to which students will validate the correctness of unknown information on the Web by relating to other websites, prior knowledge, peers or other printed materials.
2. **Authority as correctness scale (Authority):** assessing the extent to which students will examine the accuracy of unknown information in Web-based learning environments by the ‘authority’ of the websites or sources.
3. **Content as usefulness scale (Content):** measuring the extent to which students will assess the usefulness of the information viewed in Web-based learning environments by the relevancy of its content.
4. **Technical issues as usefulness scale (Technical):** assessing the extent to which students will judge the usefulness of the information viewed in Web-based learning environments by the ease of retrieval, the ease of searching or the ease of obtaining information. Therefore, their standard for evaluating Web information is more related to some technical issues.
5. **Elaboration as searching strategy scale (Elaboration):** measuring the extent to which students will have purposeful (metacognitive) thinking or integrate Web information from several websites to find the best fit that fulfills their purpose.
6. **Match as searching strategy scale (Match):** investigating the extent to which students will be eager to find only a few websites that contain the most fruitful and relevant information when they search for Web information. Their strategy is oriented towards matching the purposes of their search.
The questionnaire in this study was presented in Chinese. Some background information, such as the respondent’s
gender, major and online hours per week, was also gathered by the questionnaire. We accordingly chose six medical
universities in Taiwan and volunteer medical students from those schools to administer our paper-and-pencil
questionnaires. Students participating in this research showed interest in responding to the questionnaires.

**Result**

**Factor analysis**

To validate the instrument, we used the exploratory factor analysis by principle component method. The factor
analysis revealed that the students’ responses on the questionnaire could be grouped into six factors: ‘multiple
sources,’ ‘authority,’ ‘content,’ ‘technical,’ ‘elaboration’ and ‘match.’ These six factors were exactly the same as
those identified by Wu and Tsai (2005, 2007). These factors accounted for 68.29% of the variance. The eigenvalues
of the six factors from the principle component analysis were all larger than one, and the reliability (alpha)
coefficients for these scales were 0.76, 0.83, 0.90, 0.83, 0.68 and 0.70 respectively with the overall alpha being 0.89
as presented in Table 1. Therefore, the ICS, with six factors (scales), was deemed to be sufficiently reliable for
assessing medical students’ information commitments toward Web-based information.

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1: Multiple sources</th>
<th>Factor 2: Authority</th>
<th>Factor 3: Content</th>
<th>Factor 4: Technical</th>
<th>Factor 5: Elaboration</th>
<th>Factor 6: Match</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α=0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi. Sour. 1</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi. Sour. 2</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi. Sour. 3</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Factor 2: Authority α=0.83
| Authority 1   | 0.74                        |                     |                   |                    |                      |               |
| Authority 2   | 0.76                        |                     |                   |                    |                      |               |
| Authority 3   | 0.77                        |                     |                   |                    |                      |               |
| Authority 4   | 0.72                        |                     |                   |                    |                      |               |
| Factor 3: Content α=0.90
| Content 1     | 0.79                        |                     |                   |                    |                      |               |
| Content 2     | 0.81                        |                     |                   |                    |                      |               |
| Content 3     | 0.76                        |                     |                   |                    |                      |               |
| Content 4     | 0.78                        |                     |                   |                    |                      |               |
| Content 5     | 0.72                        |                     |                   |                    |                      |               |
| Factor 4: Technical α=0.83
| Technical 1   | 0.77                        |                     |                   |                    |                      |               |
| Technical 2   | 0.76                        |                     |                   |                    |                      |               |
| Technical 3   | 0.68                        |                     |                   |                    |                      |               |
| Technical 4   | 0.71                        |                     |                   |                    |                      |               |
| Factor 5: Elaboration α=0.68
| Elaboration 1 | 0.57                        |                     |                   |                    |                      |               |
| Elaboration 2 | 0.86                        |                     |                   |                    |                      |               |
| Elaboration 3 | 0.56                        |                     |                   |                    |                      |               |
| Factor 6: Match α=0.70
| Match 1       | 0.66                        |                     |                   |                    |                      |               |
| Match 2       | 0.83                        |                     |                   |                    |                      |               |
| Match 3       | 0.81                        |                     |                   |                    |                      |               |
| Eigen-value   | 2.18                        | 2.78                | 3.95              | 2.61               | 1.55                 | 1.96          |
| % of variance | 9.90                        | 12.62               | 17.97             | 11.84              | 7.04                 | 8.92          |

Notes: overall α = 0.89, total variance explained is 68.29%
To match the theoretical framework of the ICS, the factors are not reported in order of their extraction.
Medical students’ scores on the scales

Table 2 shows average scores and standard deviations on the six scales of the ICS for these medical students. According to Table 2, they scored highest on the ‘content’ (an average of 4.96 per item in the 1-6 Likert mode) followed by ‘elaboration’ (an average of 4.78 per item), ‘multiple sources’ (an average of 4.65 per item), ‘authority’ (an average of 4.62 per item), and ‘technical’ (an average of 4.36 per item). The lowest was ‘match’ (an average of 3.37 per item). Compared with Wu and Tsai’s results (2007), we found that the ranking of the scores on the six scales by university students in general and medical students in Taiwan was almost the same except for the ‘multiple sources’ and ‘authority’ scales.

Table 2: Students’ scores on the ICS scales (n=534)

<table>
<thead>
<tr>
<th>Scale</th>
<th># of items</th>
<th>Item mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple sources</td>
<td>3</td>
<td>4.65</td>
<td>0.82</td>
</tr>
<tr>
<td>Authority</td>
<td>4</td>
<td>4.62</td>
<td>0.86</td>
</tr>
<tr>
<td>Content</td>
<td>5</td>
<td>4.96</td>
<td>0.72</td>
</tr>
<tr>
<td>Technical</td>
<td>4</td>
<td>4.36</td>
<td>1.04</td>
</tr>
<tr>
<td>Elaboration</td>
<td>3</td>
<td>4.78</td>
<td>0.87</td>
</tr>
<tr>
<td>Match</td>
<td>3</td>
<td>3.37</td>
<td>1.17</td>
</tr>
</tbody>
</table>

Gender differences on information commitments

Table 3 shows the gender comparisons on the ICS scales identified by t-tests. It was found that compared to the female medical students, the male medical students were more oriented towards using ‘multiple sources,’ ‘authority,’ ‘technical’ and ‘elaboration’ to seek information through the Internet.

Table 3: Gender comparisons of the ICS scales (Male: n=223, Female: n=311)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Gender</th>
<th>Mean</th>
<th>S.D.</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple sources</td>
<td>Male</td>
<td>4.74</td>
<td>0.85</td>
<td>2.16*</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4.59</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td>Male</td>
<td>4.79</td>
<td>0.83</td>
<td>3.93**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4.50</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>Male</td>
<td>5.02</td>
<td>0.76</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4.91</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Technical</td>
<td>Male</td>
<td>4.50</td>
<td>1.09</td>
<td>2.67**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4.26</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Elaboration</td>
<td>Male</td>
<td>4.91</td>
<td>0.85</td>
<td>2.81**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4.70</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>Female</td>
<td>3.37</td>
<td>1.15</td>
<td>0.08</td>
</tr>
</tbody>
</table>

* p < .05, ** p < .01

The comparisons for the three advanced information commitments in this study, that is, ‘multiple sources,’ ‘content’ and ‘elaboration’ show that the male medical students attained significantly higher scores in the ‘multiple sources’ and ‘elaboration’ scales than the female students, but for the ‘content’ scale, there was no significant difference. On the other hand, for the three less sophisticated information commitments, that is, ‘authority,’ ‘technical’ and ‘match,’ the male medical students scored significantly higher for the ‘authority’ and ‘technical’ scales than the female students. In light of these findings, it can be said that male students do not always outperform female students.

The role of internet experience in information commitments

Table 4 presents the connections between student Internet usage experience and their responses on the ICS scales by using Pearson’s correlation. In this study, the average number of online hours per week was considered to represent internet experience. This method of representation is the same as that defined by Wu and Tsai (2007). We used the
following options for the respondents to represent their average online hours per week: (1) 0-10 hours, (2) 11-20 hours, (3) 21-30 hours, (4) 31-40 hours, (5) 41-50 hours, (6) 51-60 hours, (7) 61-70 hours, (8) 71-80 hours, and (9) 81-90 hours. The students’ average score on this indicator was 4.07 (suggesting an average of slightly more than 40 hours per week). The correlation analysis revealed that medical students’ internet experience played an important role in their use of all the searching strategies except for the ‘match’ scale. Students with more Internet experience tended to use the ‘multiple sources,’ ‘authority,’ ‘content’ and ‘technical’ standards to judge online information and utilize the ‘elaboration’ searching strategy while they were oriented to employ quite ‘mixed’ standards for judging online information. However, it should be noted that although the above correlations are statistically significant, they are relatively low in terms of the correlation coefficients.

Table 4: The correlation between students’ scores on the six scales of the ICS and their Internet usage experience (n=534)

<table>
<thead>
<tr>
<th>Internet usage experiences</th>
<th>Multiple sources</th>
<th>Authority</th>
<th>Content</th>
<th>Technical</th>
<th>Elaboration</th>
<th>Match</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.22***</td>
<td>0.19****</td>
<td>0.15**</td>
<td>0.11*</td>
<td>0.09*</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*p < .05, ** p < .01, *** p < .001

Differences in the information commitments of medical students and other university students

In order to better interpret medical students’ information commitments, this study further conducted a series of comparisons of the responses of the medical students and those of university students in general (n = 1220) reported in Wu and Tsai’s research (2007). The t-test comparisons of the results of these two studies are presented in Table 5.

Table 5: Comparisons of medical students and general university students’ scores on the ICS scales

<table>
<thead>
<tr>
<th>University students (Wu &amp; Tsai, 2007, n=1220) (Mean, S.D.)</th>
<th>Medical students (n=534) (Mean, S.D.)</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple sources</td>
<td>(4.65, 0.82)</td>
<td>-4.00***</td>
</tr>
<tr>
<td>Authority</td>
<td>(4.62, 0.86)</td>
<td>0.98</td>
</tr>
<tr>
<td>Content</td>
<td>(4.96, 0.72)</td>
<td>4.31***</td>
</tr>
<tr>
<td>Technical</td>
<td>(4.36, 1.04)</td>
<td>-6.50***</td>
</tr>
<tr>
<td>Elaboration</td>
<td>(4.79, 0.87)</td>
<td>2.55*</td>
</tr>
<tr>
<td>Match</td>
<td>(3.37, 1.17)</td>
<td>-7.45***</td>
</tr>
</tbody>
</table>

*p < .05   *** p < .001

Table 5 shows the comparisons for the three advanced information commitments, that is, ‘multiple sources,’ ‘content’ and ‘elaboration.’ The medical students, only on the ‘multiple sources’ scale, attained significantly higher scores than the general university students whereas they had statistically lower scores on the other two scales. On the other hand, for the three less sophisticated information commitments, the medical students scored significantly higher on the ‘technical’ and ‘match’ scales than the general university students did. For the ‘authority’ scale, there was no significant score difference between the two groups. It can thus be concluded that medical students hold less sophisticated information commitments than university students in general except for the ‘multiple sources’ and ‘authority’ scales.

In this study, we further compare the responses sorted by gender, again using t-tests. Table 6 presents the results for male students and Table 7 for female students.

Based on the analysis of the data in Table 6, for the three advanced information commitments, male medical students tended to express the usage of ‘multiple sources’ to a higher degree than general male university students, but there were no significant differences between these two groups on the ‘content’ and ‘elaboration’ scales. However, for the three less sophisticated information commitments, male medical students also possessed a stronger orientation towards less advanced information commitments than male university students in general, i.e. ‘technical’ and ‘match’ but not ‘authority.’ An analysis to compare female students shown in Table 7 revealed more critical findings.
Although there were no significant differences between female medical students and female university students in general for the commitments of ‘multiple sources’ and ‘authority,’ for the advanced information commitments such as ‘content’ and ‘elaboration,’ female medical students tended to have significantly lower scores while also having higher scores on less sophisticated information commitments such as ‘technical’ and ‘match.’

<table>
<thead>
<tr>
<th>Table 6: Comparisons of male students’ scores on the ICS scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>General university students</td>
</tr>
<tr>
<td>(Mean, S.D.)</td>
</tr>
<tr>
<td>Multiple sources</td>
</tr>
<tr>
<td>Authority</td>
</tr>
<tr>
<td>Content</td>
</tr>
<tr>
<td>Technical</td>
</tr>
<tr>
<td>Elaboration</td>
</tr>
<tr>
<td>Match</td>
</tr>
</tbody>
</table>

* p < .05  *** p < .001

<table>
<thead>
<tr>
<th>Table 7: Comparisons of female students’ scores on the ICS scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>University students (n=421)</td>
</tr>
<tr>
<td>(Mean, S.D.)</td>
</tr>
<tr>
<td>Multiple sources</td>
</tr>
<tr>
<td>Authority</td>
</tr>
<tr>
<td>Content</td>
</tr>
<tr>
<td>Technical</td>
</tr>
<tr>
<td>Elaboration</td>
</tr>
<tr>
<td>Match</td>
</tr>
</tbody>
</table>

** p < .01  *** p < .001

**Discussion and Conclusions**

In this study, we used the ICS survey developed by Wu and Tsai (2005) to explore Taiwanese medical students’ information commitments. First, by exploratory factor analysis, it was found that the ICS was sufficiently reliable for probing these students’ information commitments toward web information. We also found that medical students’ Internet experience was related to all the information commitments except for the ‘match’ scale. This study revealed that students with more Internet experience tended to utilize the ‘elaboration’ searching strategy which is more advanced while they were oriented to employing quite ‘mixed’ standards for judging the accuracy and usefulness of online information. Wu and Tsai (2005) found that university students may use both the information commitments of ‘multiple sources’ and ‘authority’ as their accuracy standard and use ‘content’ and ‘technical’ information commitments as their usefulness standard. In this study, we confirmed the same result that medical students applied quite assorted standards to seeking information through the Internet. Our results suggest that Internet experience may help students use better searching strategies (e.g., elaboration), but they may simultaneously utilize diverse sets of standards (either advanced or less sophisticated) for evaluating the accuracy and usefulness of the online information they find.

In this study, we investigated the differences between male and female medical students. We found that the male medical students used the ‘multiple sources,’ ‘authority,’ ‘technical’ and ‘elaboration’ information commitments more than the female medical students did. In other words, concerning standards of accuracy, for the two information commitments of ‘multiple sources’ and ‘authority,’ male medical students scored higher than female students. That is, on the ‘multiple sources’ scale, male medical students might have a tendency to discuss with others and then judge the information or validate it by more websites than the female students when viewing unknown information on the Internet. But, interestingly, the results of the ‘authority’ scale also indicated that the male medical students might tend to believe in the accuracy of online information if it is posted on well-known or government websites.
For the standards of usefulness, male medical students tended to use the less sophisticated ‘technical’ information commitment, but they had a greater tendency to use the advanced searching strategy, ‘elaboration,’ when they were seeking information through the Internet. That is, compared to female medical students, male medical students used a better strategy to search for information through the Internet. This finding somewhat concurred with those revealed by previous studies that female students might have unfavorable outcomes for Internet-related activities (Escoffery et al., 2005; Kim et al., 2007; Hardt & Hollis-Sawyer, 2007).

In recent years, some studies have been conducted on Taiwanese students to explore gender differences on Internet-related activities or attitudes. For example, Tsai, Lin and Tsai (2001) found that Taiwanese male high school students tend to express more favorable feelings and have lower anxiety and higher confidence about using the Internet than female students. Tsai and Lin (2004) pointed out that female Taiwanese adolescents held more pragmatic views of the Internet, and male adolescents, similar to those in Tsai, Lin and Tsai’s research (2001), expressed significantly more positive attitudes toward the Internet than females. However, they found that females tended to show higher Internet self-efficacy than males (Tsai & Lin, 2004). Chen and Tsai (2007) nevertheless found that female Taiwanese university students possessed more positive attitudes than male students toward the helpfulness and variety of the digital contents for Web learning (Chen & Tsai, 2007). Tsai (2008) found that Taiwanese male university students tended to show greater preference than female students in web-based learning environments where they could solve challenging problems, acquire cognitive apprenticeship and guidance from experts and promote epistemological development (Tsai, 2008). Similar to the findings in this study, the results of the gender differences on Internet learning or relevant attitudes did not seem very consistent.

Furthermore, this study showed that the medical students gained significantly lower scores than university students in general for two of the three advanced information commitments (scales of ‘content’ and ‘elaboration’) and only on the scale of ‘multiple sources’ did they not score significantly lower. However, for the other three less sophisticated information commitments, the medical students scored significantly higher on the scales of ‘technical’ and ‘match’ than university students in general. As for the scale of ‘authority,’ there was no significant difference between the two groups. This study reveals that medical students tend to possess less advanced information commitments than university students in general (except for their standards for judging the accuracy of web information). In particular, medical students focus more on some technical issues rather than on real content when evaluating the usefulness of online information. Also, they tend to use the ‘match’ rather than the ‘elaboration’ strategy when searching for web information. This situation is more apparent amongst female medical students.

There were no significant differences between female medical students and female university students in general in terms of the information commitments of ‘multiple sources’ and ‘authority,’ but on the scales of ‘content’ and ‘elaboration,’ which are considered to be advanced information commitments, female medical students tended to have significantly lower scores. In addition, compared to female university students in general, these female students tended to have higher scores for the less sophisticated information commitments, such as ‘technical’ and ‘match.’ These results indicate that on the usefulness standard and information searching strategy commitments, female medical students tend to have less sophisticated commitments than other female university students do when they search for information through the Internet.

In this study, we have studied the medical students’ commitments of searching for medical information through the Internet and explored gender differences. More representative samples may be needed in the future to reach more valid conclusions about these issues. Some other mediated variables, such as socio-economic status (SES) or students’ majors, may be considered for future research about the interplay between student information commitments and gender. The findings of the present study strongly suggest that medical school students need additional training, including learning how to critically assess the accuracy of online information and to develop better strategies for using web resources to seek medical information. As they implement more and more Internet-assisted instructional activities, medical educators need to find some ways or design some instructional plans to help their students, especially female students, to acquire more sophisticated information commitments. Morahan-Martin (2004) mentioned how Internet users find, evaluate and use online health information and suggested that professionals should help people identify good medical or health care information websites. We suggest that professionals should be engaged as major participants in developing proper criteria for healthcare or medical information sites for lay people (the commons or non-professional people), medical students or clinicians. Besides, medical students must learn how to evaluate the data on the Internet and need the skills to carefully use online medical information resources (Ward et al., 2002). In a recent study, we used some guided online inquiry activities
to improve a group of college students’ information commitments (Liang & Tsai, 2008), and these activities were quite successful for enhancing students’ judgments about online information. Medical educators may use similar approaches to help medical students.

Acknowledgement
The funding of this research was supported by the National Science Council, Taiwan, under grant numbers NSC 95-2511-S-011-003-MY3, 95-2511-S-011-004-MY3 and 96-2511-S-011-002-MY3.

References


Automatic Generation of Just-in-time Online Assessments from Software Design Models

Imran A. Zualkernan1, Salim Abou El-Naaj2, Maria Papadopoulos3, Budoor K. Al-Amoudi4 and Charles E. Matthews5

1American University of Sharjah, Sharjah, U.A.E. Tel: 971-50-6260243 // Fax: 971-6-5152979 // izualkernan@aus.edu
2EDS, Dubai, General Motors DWTC, P.O. Box 9233, Dubai, U.A.E. // Tel: 971-50-8115582 // Fax: +971-4-331 4102 // salim.abouelnaaj@gm.com
3International School of Choueifat, P.O. Box 15679, Dubai, U.A.E. // Tel: 971-4-2220560 // Fax: 971-4-2257955 // maria.mitchae84@gmail.com
4Thuraya Satellite Telecommunications Company, P.O. Box 29222, Sharjah, U.A.E. // Tel: 971-6-8080482 // Fax: 971-6-8828484 // b_alamoudi@thuraya.com
5Fifth Generation Systems, Ltd., Markham, Ontario, Canada // Tel: (905) 887-8522 // charles.matthews@acm.org

ABSTRACT

Computer software is pervasive in today’s society. The rate at which new versions of computer software products are released is phenomenal when compared to the release rate of new products in traditional industries such as aircraft building. This rapid rate of change can partially explain why most certifications in the software industry are generic as opposed to those in the aircraft-building industry where engineers and technicians are certified to work on a specific aircraft. For example, a software engineer may be certified on a database management system, but not on a specific implementation based on the database management system. Hence, software engineers are allowed to make critical changes to specific designs for the next release of a software product with little formal assessment of their understanding of the design. This paper presents a system that automatically generates just-in-time online assessments for judging a software engineer’s comprehension of artifacts representing software designs. The assessments thus generated are compliant with the IMS-QTI 2.1 standard. The system is based on the AXIS web-services architecture and provides a priori statistical estimates of effectiveness of each individually generated assessment.

Keywords

Design certification, Assessment generation, QTI, Web services

Introduction

New versions of software products are released every few months. Given the mission-critical nature of software products today, it is reasonable to expect that software engineers should be “certified” on a software design to ensure that they understand the design before being allowed to change it for the next release. However, this is often not the case. Most software design methodologies employ a “tell and pray” pedagogy in which the engineers are given the software product along with its design documents and are expected to pick up the design as they work; the engineers are almost never assessed to ensure that they understand the design.

In many situations, however, engineers are provided with visualization tools to help them understand software designs (Hundausen, Douglas & Stasko, 2002; Stasko, Dominique, Brown, & Price, 1988). These tools range from automatic help-file generators at the code level to browsers at the design and specification level (Confora, Cimitile, Carlini, & De Lucia, 1998; Hendrix, Cross, & Maghsoodloo, 2002; Lanza & Ducasse, 2003; Lehman, 1989; Tonella, 2003; Luqi, Berzins, & Qiao, 2004). While these tools help an engineer explore existing designs, they provide little guidance on how well an engineer understands a design.

Similarly, inspection and walkthrough methods (Anderson, Reps, & Teittelbaum, 1989; Brykczynski, 1999; Miller & Yin, 2004; Traore & Aredo, 2004) are often used as collaborative and reflective design exercises in which engineers are forced to articulate, reflect, and defend their software designs. Inspection and walkthrough methods are effective in ensuring quality of new software designs. Like visual tools, however, inspection and walkthrough methods also offer little in establishing how well an engineer understands a particular design in any meaningful way.
Using formal methods is another approach to understanding properties of software designs (Apvrille, Courtiat, Lohr, & Saqui-Sannes, 2004; Eshuis & Wieringa, 2004). However, like other approaches mentioned earlier, these methods also fail to provide an objective assessment of an engineer’s level of understanding of a design.

Obviously, it is possible to manually construct exams to judge an engineer’s level of understanding of a software design. However, every few months, typical computer software products release a newer version. Therefore, such an approach is not cost-effective: exams would need to be rewritten every time the design changes.

This paper presents an automated approach to judging a software engineer’s level of comprehension of artifacts created during software design. This approach is cost-effective because online assessments to determine an engineer’s level of understanding of a software design are automatically generated and graded; every time the design changes, assessments can be automatically re-generated.

The paper first presents a framework to formally define the problem and outline an approach. The framework is followed by a description of architecture and a prototype system based on this approach. Examples of automatically generated assessments are provided next. The paper ends with a discussion of limitations and conclusions.

Framework

The primary objective of this research is to automatically generate assessments for checking a software engineer’s level of understanding of software design artifacts or models. For example, the Unified Modeling Language (UML) (OMG-UML, 2003) currently supports thirteen such models. Each design model is typically an approximation of some aspect of the software or the world. The task of checking a model against the world is typically called validation. Similarly, the task of checking a model for internal consistency and completeness is called verification. The goal of this research is neither validation nor verification. Rather, it is to determine how closely the “mental model” or adaptation (Simon, 1983) of a software engineer matches the actual design model. Hence, if \( M \) is a design model, the purpose of an assessment is to determine how closely \( M' \), the understanding of a software engineer, matches the actual model \( M \). In other words, the goal of assessment is to approximate \( (M – M') \) or the degree to which the understanding of a software engineer differs from the actual design model.

The meaning of a software engineer’s understanding of a model, \( M \), depends on an “authentic” context of its use (Brown, Collins, & Duguid, 1989; Lave & Chaiklin, 1993). This research assumes that this understanding is inherently related to the use of model \( M \) in the context of design. For example, if a model is used to emphasize logical gaps in reasoning, an understanding of this model consists of a software engineer’s ability to perform the task of finding such logical gaps. More generally, an understanding of a particular model \( M \) is related to a set of related design tasks.

The framework has two components. The first component describes how to generate assessments from a design model, and the second describes how to determine the effectiveness of such automatically generated assessments.

Generation of assessments

On surface, the problem of automatic generation of assessments seems related to the problem of automatically generating test cases from design models (Andrews, France, Ghosh, & Craig, 2003; Bigot et al., 2004; Chow, 1978; Nebut, Fleurey, Le Traon, & Jézéquel, 2006). However, unlike test cases, which are used for validation and verification, assessments check the depth of an engineer’s comprehension of a design model. Therefore, this framework follows Bloom’s (1956) pedagogical categories as the foundation for generating assessments. Bloom provides generic categories of levels of learning for cognitive tasks, such as knowledge, comprehension, application, analysis, synthesis, and evaluation. Each successive level of learning requires a higher level of understanding. Roughly, each question in an assessment is generated using some of the following action verbs corresponding to each category:

- **Knowledge**: arrange, define, duplicate, label, list, memorize, name, order, recognize, relate, recall, repeat, reproduce, or state.
Comprehension: classify, describe, discuss, explain, express, identify, indicate, locate, recognize report, restate, review, select, or translate.

Application: apply, demonstrate, dramatize, employ, illustrate, interpret, operate, practice, schedule, sketch, solve, transcribe, use, or write.

Analysis: analyze, calculate, categorize, compare, contrast, correlate, criticize, diagram, differentiate, discriminate, distinguish, examine, experiment, prove, question, or test.

Synthesis: arrange, assemble, collect, compose, construct, create, design, develop, formulate, integrate, manage, organize, plan, prepare, propose, or set up.

Evaluation: appraise, argue, assess, attach, choose, defend, estimate, evaluate, judge, predict, rate, select, support, or value.

Therefore, for a particular design model, an application of Bloom’s categories leads to the generation of assessment questions at each of the six levels of understanding.

This research is currently focused on automatic generation of assessments for activity diagrams. Activity diagrams are a type of model commonly used in software design. An activity diagram for making a beverage is shown in Figure 1. As Figure 1 shows, an activity diagram typically captures control flow in a situation. For example, Figure 1 shows that a coffee machine has to be turned on before brewing coffee. An activity diagram also captures coordination or syncing points when activities are allowed to occur concurrently until some point in time. For example, in Figure 1, putting coffee in filter, adding water, and getting cups can happen concurrently. However, all three activities must finish before coffee can be poured. This condition is indicated by the horizontal bar in the figure. Finally, activity diagrams also include “guards” or conditions such as “coffee not found” that lead to different branching based on decision points indicated by the diamond symbol.

Figure 1. A sample activity diagram for drinking coffee (Adapted from OMG-UML, 2003)
The activity diagram shown in Figure 1 will now be used to illustrate how assessments at various levels of Bloom’s hierarchy can be generated from an arbitrary activity diagram.

One type of knowledge-level assessment is a simple recall. For example, for the activity diagram shown in Figure 1, a recall-based assessment consists of asking an engineer to indicate if “pour coffee” was one of the activities in the diagram. Another variation may be to ask a software engineer to list the various activities in the diagram.

An assessment at the comprehension level is constructed by asking an engineer to describe and explain what is in the activity diagram. For example, a software engineer may be asked if coffee can be poured (i.e., “pour coffee”) without both “brew coffee” and “get cups” finishing first. This information is embedded in the activity diagram via the horizontal bar (see Figure 1). As this example illustrates, to answer such questions correctly, an engineer needs to understand how to interpret the relationships among the various components of an activity diagram.

An assessment at the application level is constructed by asking a software engineer to apply what she knows from this activity diagram. For example, a software engineer may be asked to produce a sequence of activities (if any) that start from the “find beverage” activity and ultimately lead to the “pour coffee” activity, under the condition that no coffee was found. In order to perform correctly for this assessment, a software engineer needs to understand how to apply conditions to generate a specific path through the activity diagram.

An assessment at the analysis level consists of the ability to use facts and inferences to understand properties of the activity diagram. For example, a software engineer may be asked to describe conditions (if any) under which “get cola” and “get cups” happen concurrently. A successful performance of this assessment requires that the engineer be able to calculate the conditions under which specific paths can be taken concurrently in the activity diagram.

An assessment at the synthesis level requires building a structure or pattern from diverse elements to create new meaning in the activity diagram. For example, a software engineer may be asked to alter the activity diagram to enable the use of cups for drinking cola as well. A successful performance at this level requires the engineer to have the ability to successfully modify the existing activity diagram to satisfy additional functionality or constraints.

Finally, an assessment at the evaluation level requires judgment about the value of ideas or materials included in the activity diagram. For example, one may ask the software engineer if the activity diagram handles exceptional conditions such as the coffee machine not turning on appropriately. In general, various properties of the activity diagram such as usability, reliability, and security can be assessed at this level.

In summary, the general framework for generating assessments is based on Bloom’s taxonomy. The description and guide words at each level are used to create specific heuristics that generate questions for an arbitrary activity diagram. A typical assessment consists of a number of questions from each level of understanding.

How many questions are sufficient to ensure that an engineer’s understanding has been judged adequately? This issue is addressed by the creation of an assessment risk profile, described next.

**Creation of assessment risk profile**

The purpose of an assessment risk profile is to provide a reasonable measure of “goodness” of an automatically generated assessment for a particular activity diagram. More specifically, the purpose of an assessment risk profile is to provide an estimate of the relationship between the size of an assessment (i.e., the number of questions being asked) and the probability that an assessment of this size will catch or uncover particular types of misunderstandings.

Critical to the creation of a risk profile is the concept of a misunderstanding. Misunderstandings are commonly occurring differences between the actual model (e.g., the activity diagram) and how those differences are understood by a software engineer. For example, in the activity diagram shown in Figure 1, a software engineer may have the misunderstanding that brewing coffee comes before turning on the machine. Another misunderstanding may be that “pour coffee” does not exist at all. This research uses the Hazard Operators (HAZOP) Kim, Clark, & McDermid, 1999) scheme to generate classes of misunderstandings for activity diagrams, as shown in Table 1.
Each row in Table 1 represents a class of misunderstandings in an activity diagram. Each class of misunderstanding can have different manifestations. Table 2 shows different manifestations for each class of misunderstanding in a graphical fashion. The first column in Table 2 shows a manifestation of misunderstanding. The second column in Table 2 shows an example of a portion of an activity diagram. The third column shows how the portion of an activity diagram in the second column is misunderstood. For example, as Table 1 shows, the AS_WELL_AS class represents those misunderstandings that maintain the original intent of the activity diagram but adds additional spurious behaviors. Two manifestations of such misunderstandings (additional loop edge and additional reverse edge without removing the original edge) are shown graphically in the first and second rows of Table 2. The first row of Table 2 further shows an example of how a simple three-node activity diagram can be misunderstood by adding an additional loop edge (i.e., to node C).

<table>
<thead>
<tr>
<th>Class</th>
<th>Nature of misunderstanding</th>
<th>Manifestations of misunderstanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS_WELL_AS</td>
<td>The specific intent of the activity diagram is maintained but the misunderstanding yields additional results.</td>
<td>1. Additional loop edge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Additional reverse edge without removing the original edge</td>
</tr>
<tr>
<td>PART_OF</td>
<td>Only some of the intention in the activity diagram is achieved.</td>
<td>3. Missing an edge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Missing an activity (and reconnected edges)</td>
</tr>
<tr>
<td>REVERSE</td>
<td>Reverse flow — flow in the activity diagram is in the wrong direction.</td>
<td>5. Reversed edge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Two consecutive activities swapped</td>
</tr>
<tr>
<td>OTHER_THAN</td>
<td>A result other than initial intention is achieved while maintaining the structure of the activity diagram.</td>
<td>7. One activity substituted for another</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. An additional activity between two activities</td>
</tr>
</tbody>
</table>

Another general class of misunderstandings in Table 1 is PART_OF. Manifestations of this class of misunderstandings are missing edges or activities: the software engineer believes that the activity does not include a specific activity or an edge. A third class of misunderstandings in Table 1 is represented by the REVERSE class. In the case of edges, the REVERSE class translates into the belief by a software engineer that an edge is reversed. In the case of an activity, REVERSE misunderstanding may exist if the engineer believes that two contiguous activities are actually swapped. Finally, the OTHER_THAN misunderstandings are those cases in which the original structure of the activity diagram is believed to be about the same but additional content is substituted. For example, an engineer may have the right structure of the activity diagram, but may substitute or confuse one activity for another. A more extreme case may be the inclusion of an additional activity between two existing activities.

The “goodness” of an automatically generated assessment can be judged by the assessment’s ability to catch a particular type of misunderstanding. This is done by using a variant of mutation testing (Boland, Singh, & Cukic, 2003; DeMillo, Lipton, & Sayward, 1978; Howden, 1982). Mutation testing has been used to evaluate test case suites (i.e., a set of test cases). A mutant (or variation) of a computer program is generated by deliberately applying a mutation operator to the program. Each mutation operator represents a fault (e.g., a misspelled variable). The mutated program is then tested using the test suite. If none of the test cases in the test suite fail, then the test suite has failed to kill the mutant. The failure to kill the mutant means that the specific test suite is not suitable for catching the fault that was introduced.

Since an assessment consists of a set of questions for a software engineer, the assessment acts much like a test suite. Similarly, the activity diagram is analogous to a computer program, and a misunderstanding is analogous to a fault. Therefore, a mechanism similar to mutation testing can be used to evaluate the effectiveness of the assessments thus generated.

Let \( a \in A \) represent an arbitrary activity diagram. A mutation operator \( \mu \in M \) is applied to the activity diagram \( a \) to generates a set of mutated activity diagrams \( A' \) where \( A' = \{ a' \mid a' \leftarrow \mu(a) \} \). The list of mutation operators (i.e.,
$M$ is based on the misunderstandings shown in Table 2. For example, one mutation operator adds an additional loop edge to an activity diagram, another mutation operator reverses an edge, and so on.

Let $Q = \{q | q \leftarrow BAG(a)\}$ represent an assessment or a set of automatically generated questions using the approach described in the previous section. (The approach is called Bloom’s assessment generator [BAG].) Each $q \in Q$ represents one question for the software engineer based on Bloom’s taxonomy.

Finally, for a particular $q \in Q$, $a \in A$ and $a' \in A'$, let $ORACLE(q,a,a')$ represent an oracle that returns true if the question $q \in Q$ produces the same answer for both $a$ and $a'$, and false if it does not. For example, suppose the question “Is ‘add water’ an activity in the activity diagram?” was generated from the activity diagram shown in Figure 1. The correct answer for this question, based on the original activity diagram, is obviously “yes.” However, suppose a misunderstanding were introduced in the original activity diagram by removing the activity “add water.” This is done by applying the mutation operator (4) from Table 2. The resulting mutated activity diagram would be identical to the original diagram except that the “add water” activity and associated edges would be missing. If one were to answer the original question using the mutated diagram, the answer would be “no.” This means that the oracle answers false in this case because the answer to the question in the original and the mutated diagrams are different. A false answer from the oracle means that a misunderstanding was caught by the question.

<table>
<thead>
<tr>
<th>Manifestation</th>
<th>Actual</th>
<th>Misunderstood as</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Additional loop edge</td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
<tr>
<td>2. Additional reverse edge without removing the original edge</td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
<tr>
<td>3. Missing an edge</td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td>4. Missing an activity and reconnected edges</td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
<tr>
<td>5. Reversed edge</td>
<td><img src="image9" alt="Diagram" /></td>
<td><img src="image10" alt="Diagram" /></td>
</tr>
</tbody>
</table>
6. Two consecutive activities swapped

7. One activity substituted for another

8. Additional activity between two activities

An algorithm to generate a risk profile for a specific activity diagram $a \in A$ takes a set $P = \{a, M, k, n\}$ as input, in which $k$ represents the number of questions in the assessment being generated, $n$ represents the number of replication representing statistical accuracy, and $M$ is a set of mutation operators. The algorithm carries out binomial experiments that produce an estimated probability that an assessment with $k$ questions generated from activity diagram $a$ will catch a particular type of misunderstanding. An outline of the algorithm for the system is shown below.

Given $P = \{a, M, k, n\}$
For each $\mu \in M$ do // for all mutation operators
    $A' = \{a' | a' \leftarrow \mu(a)\}$ // generate the set of mutated activity diagrams for $a$
For each $a \in A'$ do // for all mutations of $a$
    $success(a') \leftarrow 1$
For $i = 1, n$ do // repeat the experiment $n$ times
    // randomly generate an assessment with $k$ questions
    $Q = \{q | q \leftarrow BAG(a)\}$ where $|Q| = k$
For each $q \in Q$ do
    If $(ORACLE(a, a', q) \equiv true)$ // the mutated and original gave
    // the same answer
    then $success(a') \leftarrow 0$ // one misunderstanding was not caught
\[ p(\mu, a, k) \leftarrow \frac{\sum_{\text{success}(i)} \left( n \times |A| \right)}{n \times |A|} \]

The algorithm returns the probability $p(\mu, a, k)$, that for a specific activity diagram $a$, an assessment with $k$ questions will catch a particular misunderstanding injected by a mutation operator $\mu$. Since each application of the oracle represents a binomial experiment, the statistical bounds on the probability $p$ are given by
\[ p \pm z_{\alpha/2} \sqrt{p(1-p) / (n \times |A|)} \]

where \( z \) represents the normalized score while \( \alpha \) represents the Type I error rate.

In summary, the risk profile for each activity diagram is generated by introducing known misunderstandings repeatedly in the same activity diagram, randomly generating assessments of a particular size and determining how many of these assessments are able to catch the introduced misunderstandings. This procedure leads to a statistical profile of how well the technique works for a specific activity diagram in terms of unearthing various types of misunderstandings.

**System architecture**

A prototype system based on the framework presented earlier was constructed to automatically generate online assessments from arbitrary activity diagrams. As shown in Figure 2, the system architecture is based on the Apache Axis framework for building web services (Axis, 2008). The system relies on Apache Tomcat (Tomcat, 2008) and uses MySQL (MySQL, 2008) as the back-end database. The system has one primary server that provides assessment generation services by optionally using distributed clients to calculate the risk profile for an arbitrary activity diagram. The assessment generation is not computationally expensive because it relies on known graph algorithms and pattern matching. The calculation of risk profile, however, is computationally expensive because for each profile all possible mutations of a particular type (say, reverse edge) are applied to an activity diagram, and this process is replicated depending on the statistical accuracy desired. Distributed clients essentially help tackle this portion of the computational complexity. The server supports two types of clients. Submit clients (SC) are used to submit activity diagrams, and processing clients (PC) are used to help generate the risk profile for a particular activity diagram. Each of these is described below.

**Server**

The server supports two SOAP (Soap, 2008) interfaces. The first interface is a submit activity diagram interface (S). This interface allows any SOAP client to submit an activity diagram to generate an assessment. In addition to submitting an activity diagram, the client also specifies the number of questions in the assessment (\( k \)), the mutation operators to be used, and a sample size (\( n \)) for the risk profile. The server, in turn, generates and returns an assessment and a risk profile to the client. The second interface is a processing client interface (P) that distributes and runs mutations for the risk profile and returns the results to the server.
The server has three primary components: the Bloom assessment generator (BAG), the work distribution layer (WDL), and the results tabulation layer (RTL). A SOAP client connects to the submit interface (S) of the server to provide a job to execute. A typical job consists of an activity diagram in the form of an XMI file (OMG-XMI, 2006), the mutation operators to apply, the number of questions in each assessment, and the sample size. The BAG takes the XMI file representing an activity diagram as input and generates an assessment for the activity diagram as an XML file based on the IMS QTI format (IMS-QTI, 2006).

```plaintext
Generate_order_know_13(activity_diagram a)
{
    a_u = randomly select an activity from activity_diagram a
    a_v = randomly select another activity from activity_diagram a

    \{a_u, a_v\} = Generate m activities that lie on path between a_u and a_v

    Generate the QTI 2.1 code for the following question template:

    "Arrange a_u, a_v, \ldots, a_n performed between activities a_u and a_v in the correct sequential order."

    with the correct answer being the order 1,\ldots,n.
}
```

*Figure 3.* A knowledge-level heuristic generating a multiple-choice question

```plaintext
Generate_t/f_comp_3(activity_diagram a)
{
    Randomly jump to first: or second:

    first:
    \{a_i, a_j\} = Find a random pair of activities a_i and a_j that share a join in a correct_answer = true;
goto generate:

    second:
    \{a_i, a_j\} = Find two activities a_i and a_j that are strictly sequential (a_i follows a_j)
correct_answer = false;

generate:

    Generate the QTI 2.1 code for the following question template:

    "Are the activities a_i and a_j potentially concurrent?"

    with correct answer being correct_answer;
}
```

*Figure 4.* A comprehension-level heuristic for generating a true/false question
BAG is implemented using Java, SWI-Prolog (Clocksin & Mellish 1994; SWI-Prolog, 2008), and JPL Java APIs (http://sourceforge.net/projects/jpl). BAG consists of a collection of heuristics that are organized according to Bloom’s levels. Each level incorporates a number of pedagogical heuristics. From a learning perspective, a software engineer needs to learn the sequencing and conditions under which particular activities can occur in an activity diagram. This is reflected in most heuristics. These heuristics are encoded as rules in the Prolog programming language. Each heuristic is parameterized with respect to the various components of an activity diagram. Figure 3 shows the pseudo-code for one such heuristic at the knowledge level. This heuristic is able to generate questions that ask a software engineer to order a set of activities between two arbitrarily selected activities \((a_x, a_y)\) for a particular activity diagram.

Figure 4 shows an example of a heuristic that generates a true/false question asking the user to select activities that can be potentially concurrent. As the figure shows, the heuristic randomly generates both positive (true) and negative (false) questions.

Finally, Figure 5 shows an example of a heuristic at the synthesis level. This heuristic taps into a software engineer’s ability to effect changes to the existing activity diagram and hence is at the synthesis level. In order to be able to answer questions generated from this heuristic, the software engineer needs to work through the consequences of changes in the activity diagram.

```prolog
Generate_t/f_syn_12(activity a)
{
    a_x = randomly select an activity from activity diagram a.
    a_y = randomly select an activity that follows activity a_x.
    a_z = randomly select another activity that follows activity a_x.

    Generate all paths between a_x and a_z after deleting an edge between a_x and a_y. If there is at least one path then correct_answer = true. Otherwise, the correct_answer = false.

    Generate the QTI 2.1 code for the following question template:

    “If an edge between activity a_x and activity a_y is deleted, will the user still be able to reach activity a_z?”

    with the correct answer being correct_answer.
}
```

Figure 5. A synthesis-level heuristic for generating a true/false question

The WDL takes the activity diagram, the generated assessment, and the sample size representing statistical accuracy and splits and assigns various mutation tasks for generating the risk profile to different processing clients. Finally, the results tabulation layer (RTL) consolidates the results received from various processing clients.

Once an assessment is generated, the WDL divides and assigns the job to various processing clients to generate the risk profile for the job. For example, the sample of jobs submitted to the server shown in Table 3 for two activity diagrams will be distributed as tasks to various processing clients as shown in Table 4.
Table 3. Sample jobs submitted to the server

<table>
<thead>
<tr>
<th>Job</th>
<th>$M$</th>
<th>$n$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>{1, 2}</td>
<td>100</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>{2, 3}</td>
<td>50</td>
<td>10</td>
</tr>
</tbody>
</table>

$M$ is a set of mutation operators

$n$ is the number of replication representing statistical accuracy

$k$ is the number of questions in the assessment being generated

Table 4. Mapping of jobs to tasks

<table>
<thead>
<tr>
<th>Task</th>
<th>Job</th>
<th>$\mu$</th>
<th>$n$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>3</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>3</td>
<td>25</td>
<td>10</td>
</tr>
</tbody>
</table>

$\mu$ is a mutation operator

$n$ is the number of replications representing statistical accuracy

$k$ is the number of questions in the assessment being generated

The server waits for processing clients to attempt connection and assigns a unique identifier to each client. Once the client is connected, the server sends the parameters for the task, along with the required files, to the processing client. When the client sends back the results for the task it received, the server records these results in its database and sets the task to be completed.

**Submitting clients**

Since the server supports the SOAP protocol, clients can be written in any programming language that supports the SOAP protocol. These clients connect to the server through the $S$ interface as shown in Figure 2. Figure 6 shows one such SOAP client (implemented in Java) that allows the user to select the mutation operators and submit an activity diagram. The submitting client is able to watch the progress on its job via a status bar.

**Processing clients**

Client nodes that implement the algorithm described earlier have three layers as shown in Figure 2. These clients connect to the server through the processing client (P) interface. The mutation analysis framework layer (MAFL) provides services related to creating mutations and applying the ORACLE (OR) to determine the results. In addition, each client includes BAG to randomly create an assessment of a particular size (number of questions) while the ORACLE determines whether a question yields the same results for the normal and the mutated activity diagrams.

Each processing client initiates communications with the server. When the client has resources, it continuously attempts to connect to the server in order to receive a unique identifier. Once the connection is established, the client requests the task from the server and, upon receiving the job, the client starts mutation testing based on the parameters it has received. After completion, the client calls the return result method to give back the results to the server. After dispatching the results, the processing client probes the server for any additional tasks to process. The MAFL is organized according to the HAZOP classification described earlier. Currently, the MAFL includes all the misunderstandings in Table 2 implemented as mutation operators.
Figure 6. A sample submit client

Evaluation

The system currently contains 115 rules for generating questions at various levels of the Bloom’s taxonomy. Figures 7 and 8 show sample questions from an assessment that was automatically generated from the activity diagram shown in Figure 1.

Figure 7. A knowledge-level sample question for the drink-coffee activity diagram
Figure 9 shows how the results of the assessment are automatically scored and presented to a software engineer through a commercial learning management system supporting the IMS QTI assessment format.

The risk profile for the drink-coffee activity diagram shown in Figure 1 (labeled Activity B) is shown in Figure 10. Figure 10 shows that as the number of questions in an assessment increases, so does the probability of catching each type of misunderstanding. However, based on the current set of rules, some misunderstandings are easier to catch than others. For example, as Figure 10 shows, the reversed-edge misunderstanding, where an engineer believes that an edge is pointing in the wrong direction, is fairly easy to catch because the number of questions in an assessment is increased beyond five. On the other hand, misunderstandings such as an additional reversed edge, where an engineer believes that an additional reversed edge exists between two activities, are difficult to catch; the probability of catching such misunderstandings for this specific activity diagram is less than 30 percent even with an assessment containing twenty questions. In addition, as the number of questions is doubled from ten to twenty, the probability of catching this misunderstanding only goes up by about 10 percent.
The risk profile for the activity diagram shown in Figure 10 also provides additional useful information. For example, if an engineer generates an assessment with ten questions, the risk profile shows that there is about a 50 percent probability ($p = 0.5044$ with a 95 percent confidence interval of $+/-0.0224$) that a misunderstanding like the additional loop edge will be missed by the assessment.

Figure 10. Risk profile for the drink-coffee activity diagram

Figure 11. Risk profile for an activity diagram with eight activities
Figure 11 shows the risk profile for another activity diagram (Activity Diagram A) that is much smaller (only eight activities) and simpler than the Activity Diagram B. Figure 12 shows that for Activity Diagram A, the additional reversed-edge misunderstanding is also the most difficult (lowest probability) to catch. However, the probability of catching this misunderstanding with a ten-question assessment is much higher ($p = 0.36$ with a 95 percent confidence interval of $+/-0.033$) than the probability for Activity Diagram B ($p = 0.18$ with a 95 percent confidence interval of $+/-0.017$).

Each activity diagram has a unique risk profile that is not necessarily tied to its size. For example, Figure 12 shows the risk profile for an activity diagram (Activity Diagram C) with 24 activities. The probability of catching a reversed-edge misunderstanding for this activity diagram is actually higher ($p = 0.28$ with a 95 percent confidence interval of $+/-0.18$) than a smaller activity diagram (Activity Diagram B, for example).

In addition to the results shown above, the system presented in this paper has been successfully tested on one hundred arbitrary activity diagrams collected from published sources (mean number of activities/activity diagram = 15.56, $SD = 8.22$; mean number of edges/activity diagram = 13.65, $SD = 8.26$). Risk profiles were also generated for these diagrams by varying the size of the assessments from 1 to 20 questions in increments of five questions. Thirty Pentium IV machines were used as processing clients. As expected, each activity diagram resulted in a unique risk profile.

The methodology and the tools presented in this paper were also applied to a portion of a mission-critical software system called the Gas Compliance System (or GCS Energy) (GCS, 2008). GCS Energy is an integrated suite of software modules designed to assist a natural gas utility with compliance tracking and auditing. In the United States, the Department of Transportation’s Office of Pipeline Safety (OPS) mandates periodic inspection, testing, and maintenance tasks for gas pipeline systems through DOT Part 192 — Transportation of Natural or Other Gas by Pipeline: Minimum Federal Safety Standards. This federal standard requires that pipeline operators perform compliance tasks at thousands of sites, covering hundreds of miles. Upon completion, these tasks must be recorded, stored, and made available for review in the event of an OPS audit. Failing an audit can result in the leverage of large fines against a natural gas utility.
Based on principles found in the American Gas Association’s GPTC Guide, GCS Energy’s system manages compliance tracking for various facilities within a natural gas utility company. The various subsystem modules that track compliance activities include atmospheric corrosion, corrosion tracking, leak survey, leak tracking, pressure control stations, exposed pipe exam, pipeline patrol, and valve tracking.

One portion of the valve tracking system manages inspections. Pipeline operators must inspect every installed valve within a periodic time frame as determined by the DOT Part 192 regulations. Various valve parameters determine not only the inspection frequency but also the inspection questions and whether the questions are mandatory or optional. Figure 13 displays some of these parameters in the valve detail window.

The corresponding activity diagram that models the valve inspection functionality for GCS Energy is shown in Figure 14.

![Figure 13. Sample screen for entering valve parameters in GCS](image)

Some of the questions automatically generated by the system described here for the activity diagram shown in Figure 15 are discussed below.

Q1: Starting from the initial state, can an operator do “Return to previous window” without taking the decision ‘valid required attr. values’?

The correct answer to this question is “yes.” The software engineer needs to know that an operator can choose to cancel the inspection task and return to the previous window. This operation skips any validation checks because the user is discarding the inspection task. An incorrect answer to this question exposes a serious lack of understanding on the part of the software engineer who does not understand that a user is able to cancel the inspection task and go back to the previous window before the required attribute values are validated.

Q2: What necessary action must be performed in order to do “Update database”?

A) Record completed task data  
B) Display Message: Required task question not answered  
C) Return to previous window

The correct answer to this question is A. Again, the software engineer needs to know that C cannot occur before A and that receiving the error message in B means that the database cannot be updated. If a software engineer fails to
answer this question, he does not understand that all inspection questions are answered before the database system is updated. In this context, a software engineer who selects C as the possible answer shows a complete lack of understanding of the system because returning to the previous window can only be done either if the inspection is cancelled or if all the requisite inspection tasks have been completed.

Figure 14. An activity diagram for the valve inspection functionality in GCS

Q3: After the condition "completion date < in_service date", which of the following are possible immediate successors?

A) Display Message: Inspection task record cannot be performed prior to equipment install date
B) Update task record with new due and compliance dates
C) Update database

The correct answer to this question is A; the software engineer needs to know that A is the result if the condition test fails. In addition, the software engineer also needs to know that additional questions must be answered before the new due and compliance dates can be calculated or the database updated. A software engineer with the requisite knowledge of the design also knows that “completion date < in_service date” occurs before these other decision nodes. Therefore, A is the only correct answer.
The three questions discussed above show that questions generated by the assessment system probe for a deep understanding of a software engineer’s knowledge of the software product design, and a failure of such understanding can lead to potentially catastrophic consequences.

**Discussion and limitations**

The prototype system constructed shows that it is possible to generate useful assessments automatically from activity diagrams. In addition, a unique risk profile is provided for each activity diagram.

As is evident, the assessments at the lower levels of Bloom’s taxonomy are fairly context free and can be derived easily from the syntactic structure of activity diagrams. However, since activity diagrams are semi-formal in nature, a formal model should lead to a more robust set of questions. As one progresses to higher levels of Bloom’s taxonomy, the task dimension also becomes more important in one’s ability to generate the right rules. A model of the task can help in the generation of such questions.

The mutation operators used in this research are based on the HAZOP model. Additional operators can be derived from bug checklists (Thelin, Runeson, & Wohlin, 2003). Currently, the risk profile assumes a uniform distribution while generating the assessment questions. However, once a correspondence between the types of misunderstandings and their frequency in a software engineering environment is established, the assessment generation process can incorporate this information by increasing the proportion of questions that target frequently occurring misunderstandings in a particular environment.

Finally, the results presented in this paper are being extended to include the other twelve models of UML.

**Conclusion**

As designed artifacts like computer software become more complex and life cycles become shorter, the assessment techniques of the type presented here will become a necessary part of any design cycle. This paper describes a distributed system that automatically generates standards-based, tool-independent, and just-in-time online assessments from arbitrary activity diagrams. A key feature of this system is that, in addition to automatically generating an online assessment, the system provides a unique a priori risk profile for each activity diagram.

The results presented in this paper are limited to one specific model (i.e., the activity diagram) within one methodology (i.e., UML) in the context of one engineering discipline (i.e., software engineering). In addition, most assessment heuristics are at lower levels of Bloom’s hierarchy. However, the results presented here have a potential application in any branch of engineering that constructs formal or symbolic by-products or artifacts as a natural design activity. For example, CAD/CAM systems for mechanical and civil engineering design are obvious candidates. Another significant area for application of this approach is process design. Increasingly, businesses’ processes are not only being formalized but are in continuous flux as businesses continuously respond to changes in their environment. This flux creates a similar problem where the changes in a business’s process are to be conveyed to all individuals that play a role or come in contact with the process. Again, rather than showing an individual a “picture” of a changed business process, an approach similar to the one presented in this paper can actually generate assessments to ensure that process changes have actually been understood.

Using mutation-analysis to generate an a priori estimate of risks associated with a particular assessment is also a promising side benefit of this approach. However, the effectiveness of this method relies on the availability of information on the types of misunderstandings that typically occur in specific contexts. The approach has the potential to be generalized to other disciplines as well. For example, one can imagine a common set of misunderstandings of a construction blueprint in the context of civil engineering: a missing door, the wrong type of HVAC pipe, etcetera. Once a taxonomy of such misunderstandings is established in a field, it can be used to create mutations of an existing design artifact (such as a blueprint) to judge the effectiveness of automatically generated assessments.
In hindsight, the difficulty of constructing heuristics to generate Bloom’s higher level assessments is not surprising; it is difficult enough to construct such assessments by hand. However, in this research it was possible to arrive at some rudimentary heuristics at Bloom’s higher levels, such as synthesis. Constructing high-level heuristics presents an interesting challenge for next phase of this research. Current attempts at constructing such higher-level heuristics suggest that ultimately domain-specific task models may be required to construct heuristics at higher level of Bloom’s taxonomy. For example, a sophisticated heuristic at the synthesis level will have to generate synthesis tasks for a software engineer and automatically mark the engineer’s performance on such tasks. The tasks need to go beyond the current synthesis heuristics that require an engineer to work through the consequences of various types of changes to a design artifact. Asking a software engineer to propose a novel solution in the context of an existing design artifact and automatically generating assessments to judge the feasibility of such solutions are non-trivial tasks that will likely require a sophisticated task and domain model.

In summary, this paper represents a first set of experiments in automatically generating just-in-time assessments in the limited domain of software engineering. This approach can be applied to a large number of engineering and business contexts. The ultimate success of such an approach, however, will depend on one’s ability to generate specific heuristics for each domain.

Acknowledgements

This research was supported in part by a faculty research grant from the IBM Corporation.

References


Main Barriers and Possible Enablers of ICTs Integration into Pre-service Teacher Education Programs

Yuksel Goktas
Department of Computer Education & Instructional Technology, Ataturk University, 25240 Erzurum, Turkey // Tel: 90-546-263-5774; Fax: 90-442-236-0955 // yukselgoktas@atauni.edu.tr; yuksel.goktas@hotmail.com

Soner Yildirim and Zahide Yildirim
Department of Computer Education & Instructional Technology, Middle East Technical University, 06531 Ankara, Turkey // Tel: 90-312-210-4057 Fax: 90-312-210-7986 // soner@metu.edu.tr; zahidey@metu.edu.tr

ABSTRACT
The purpose of this study is to investigate the main barriers and possible enablers for integrating information and communication technologies (ICTs) in Turkey’s pre-service teacher education programs. The data were collected by means of questionnaires from 53 deans in schools of teacher education (STE), 111 teacher educators, and 1,330 prospective teachers, and additionally from interviews of six teacher educators and six prospective teachers. The findings indicate that the majority of the stakeholders believe that lack of in-service training, lack of appropriate software and materials, and lack of hardware are the main barriers for integrating ICTs in pre-service teacher education programs. There was also agreement on the possible enablers; “having technology plans” was the strategy most strongly agreed upon in that category. Based on these findings, we propose the use of several strategies that should enhance successful ICTs integration.

Keywords
ICT barriers, ICT obstacles, Technology integration problems, ICT enablers

Introduction
Increasing the quality of teaching and learning has been a seemingly important concern for education. Since the beginning of this century, education has faced a variety of social, cultural, economical, and technical challenges. As the study and practice of facilitating learning and improving performance (Januszewski & Molenda, 2008), the field of educational technology attempts to overcome challenges by developing new approaches and frameworks. In this context, information and communication technologies (ICTs) represent a new approach for enhancing the dissemination of information and helping to meet these challenges. ICTs comprise the use of at least a computer and the Internet as well as computer hardware and software, networks, and a host of devices that convert information (text, images, sounds, and motion) into general digital formats (Lever-Duffy, McDonald, & Mizell, 2003; USDE, 2000; ISTE, 1999).

A predetermined process is important for the integration of ICTs in the classroom, curriculum, school management, library, and any educational setting. Integration of ICTs enhances the quality of education by helping teachers to do their job and by helping students to learn more effectively. In these contexts, teachers’ shifting role in the 21st century involves an essential mission, which is to be the frontier for applying technological innovations to the teaching and learning process. At this point, necessary skills and the level of future teachers’ readiness are key factors in implementing new ICTs. Consequently, schools of teacher education play a crucial role in preparing future teachers to become proficient in the integration of ICTs into the curriculum. They need to help prospective teachers understand how ICTs can be used to teach content in rich and meaningful ways (ISTE, 1999; Keating & Evans, 2001; Roblyer & Edwards, 2000).

On the other hand, integration of ICTs into pre-service teacher education is critical to integrating ICTs in K–12 schools. Despite the huge investment of financial and human resources, pre-service teacher education programs do not currently provide prospective teachers with the necessary skills, competencies, and experiences to prepare them to use ICTs effectively in their future profession (Duran, 2000; Moursund & Bielefeldt, 1999; Bullock, 2004; Mehlinger & Powers, 2002). In light of the above-mentioned literature, it is obvious that integration of ICTs into pre-service teacher education is influenced by many barriers.
Bromme, Hesse, and Spada (2005) said of a barrier: “It comes from psychological research on problem solving and creativity. There it refers to the gap between an initial and end state. In other words, barriers are challenges which have to be overcome in order to attain a goal” (p. 1). The authors also stated that the localization of difficulties always depends on theoretically based assumptions concerning the nature of barriers. Working with ICTs is often difficult, simply because ICTs are new, and because individual and social routines have to be established in using them. Additionally, the use of ICTs is complicated because it involves not only the use of alternative tools for dealing with old, conventional problems but also expectations that these technologies will help in meeting new challenges.

A variety of action plans have been developed to effectively integrate ICTs in pre-service teacher education programs, but many barriers still exist in practice. To facilitate these plans, barriers need to be identified so that they may be overcome. ICTs integration in pre-service teacher education programs continues to be a challenge all around the world. To ease this struggle, Ertmer, Addison, Lane, Ross, and Woods (1999) noted that “when educators and researchers look for reasons why teachers are struggling to use ICTs effectively, it may be important to look at what they have (in terms of beliefs and practices) in addition to what they do not have (in terms of equipment)” (p. 68). Main barriers can be identified when incorporating ICTs in education. Toward this end, Ertmer (1999) classified barriers as falling into two primary categories: extrinsic (first order) and intrinsic (second order). Extrinsic barriers include lack of resources, inadequate training, insufficient technical support, and lack of time; intrinsic barriers include teachers’/instructors’ beliefs, visions concerning technology integration, and views about teaching, learning, and knowledge (p. 51–52).

To better prepare prospective teachers and to overcome these barriers, enablers are required. Ertmer, Ottenbreit-Leftwich, and York (2006–2007) also classified enablers, like barriers, as either extrinsic or intrinsic. For instance, access to hardware, quality software, the Internet, and technical, administrative, and peer support might be viewed as extrinsic enablers, whereas personal beliefs, previous success with technology, and self-efficacy might be viewed as intrinsic enablers (p. 55).

| Table 1. A summary list of the barriers affecting ICTs integration in pre-service teacher education programs |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Lack of in-service training | √ | – | – | – | – | – | – |
| Lack of appropriate software/materials | – | – | √ | – | – | – | – |
| Lack of basic knowledge/skills for ICTs | – | – | – | √ | √ | √ | √ |
| Lack of hardware | √ | – | √ | √ | – | – | – |
| Lack of knowledge/skills for ICT integration | – | – | √ | – | – | – | – |
| Lack of technical support | – | – | √ | – | – | – | – |
| Lack of appropriate course content and instructional programs | – | – | – | √ | – | – | – |
| Lack of time | √ | – | – | – | √ | – | – |
| Lack of appropriate administrative support | – | – | – | √ | – | √ | – |

Prior studies have identified numerous barriers and enablers that affect ICT integration in pre-service teacher education programs. Tables 1 and 2 provide summary lists of common barriers and enablers, respectively. There is currently a need, however, for further exploration of the main barriers and potential enablers that are important in the schools of teacher education (STE) in Turkey. Therefore, the purpose of this study is to reveal which main barriers are commonly encountered and which enablers are potentially the most useful for contemporary ICT integration in pre-service teacher education programs, specifically in Turkey. Consequently, this study addresses the following research questions:
1. What are the main barriers faced while integrating ICTs in pre-service teacher education programs according to deans, teacher educators, and prospective teachers?

2. What are the possible enablers for integrating ICTs in pre-service teacher education programs according to deans, teacher educators, and prospective teachers?

### Table 2. A summary list of the enablers affecting ICT integration in pre-service teacher education programs

<table>
<thead>
<tr>
<th>Method</th>
<th>Overview</th>
<th>Participants</th>
<th>Method</th>
<th>Overview</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In this study, the researchers used both quantitative and qualitative research approaches in the data collection and analysis processes. Two questionnaires were used for the quantitative data collection from the deans and teacher educators. Open-ended questions from the questionnaires and interviews were used for the qualitative data collection from the deans, teacher educators, and prospective teachers. By using different sources and approaches, the researchers intended to strengthen the validity of the results.</td>
<td>There are 63 public and private STE that train teachers for primary and secondary education in Turkey (as of Spring 2005). The deans of these schools constitute the first population in our research. This entire group was surveyed in March and April 2005. Follow-up questionnaires were sent in May and June 2005 to those deans who did not respond to the first query. In total, 51 deans responded to the deans’ questionnaire, yielding a return rate of 81 percent.</td>
<td>Initially, the teacher educators and prospective teachers were clustered into 12 statistical regions using NUTS (Nomenclature of units for territorial statistics) level 1, so as to be representative of the population. After that, 18 STE (including at least one school from each region) were selected by the convenience sampling method. Thus, 223 teacher educators and 2,116 prospective teachers were selected from these schools in May 2005 and invited to participate in the study by completing the questionnaire. Follow-up questionnaires were sent in June and July 2005 to those who did not respond to the first query. Overall, 111 teacher educators and 1,330 prospective teachers responded to the questionnaire, yielding return rates of 49.8 percent and 62.9 percent.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
respectively. The researchers also collected in-depth data via interviews, which were used to support the quantitative data. The interviews of certain participants were done by selecting six teacher educators from three STE and six prospective teachers from two STE in the capital city (Ankara) of Turkey. First, the capital city and STE were selected by the convenience sampling method. Then, six teacher educators and six prospective teachers were chosen by means of a purposeful sampling approach using the criterion technique. The criteria used for the selection of the teacher educators’ group were as follows: 1) instructs ICT-related courses; and 2) has at least three years of teaching experience in STE. The criterion used for the selection of the prospective teachers’ group was that the candidate must have taken ICT-related courses before the spring semester of 2005. All of the participants in this study have been assigned pseudonyms (e.g., A, B, C) to protect their identities.

Data collection instruments

The data were collected by means of five different instruments: three questionnaires and two interview guides. The prospective teachers’ questionnaire was developed by Tinmaz (2004). The deans’ and teacher educators’ questionnaires and both interview guides were developed specifically for this study in 2004, following a review of similar instruments in the literature (Baron & Goldman, 1994; Queitzsch, 1997; Roblyer, 1994; SEIRTEC, 1998; Smith, 2002; Topp, Mortensen, & Grandgenett, 1995; Vagle & College, 1995). Generally, items in all of them were grouped around four major topics: 1) demographic and institutional information, 2) main barriers, 3) possible enablers, and 4) comments/proposals. The prospective teachers’ questionnaire consisted of demographics and open-ended responses. However, excluding demographics and open-ended items, two questionnaires consisted of nine barrier and seven enabler items formatted as Likert-type statements to which participants responded with a rating of 1–5 (5 indicating Strongly Agree, 4 indicating Agree, 3 indicating Neutral, 2 indicating Disagree, and 1 indicating Strongly Disagree).

To determine the content validity of the instruments, we used peer and expert juries as well as a language expert. For this purpose, after the questionnaires and interview guides were developed, each instrument was examined first by three PhD candidates and then by four experts. After revisions were made, the teacher educators’ questionnaire was piloted for reliability with a group of 64 teacher educators; an internal consistency method was used, which yielded a reliability coefficient of 0.87 (Cronbach’s Alpha coefficient; see Cronbach, 1990). For the prospective teachers’ questionnaire, the pilot test reliability coefficient was 0.86 (provided by Tinmaz, 2004). A pilot test could not be conducted for the deans’ questionnaire because of the group’s characteristics. At the end of the data collection process, each set of items was re-tested, yielding reliability coefficients of 0.91, 0.97, and 0.91 for the deans’, teacher educators’, and prospective teachers’ questionnaires, respectively. All of these values are higher than the 0.80 criterion, which is regarded as internally reliable (Bryman & Cramer, 1997).

The interview guides were piloted with the help of two prospective teachers and one teacher educator to determine if the interview procedures were acceptable and if any additional interview questions were needed to answer the research questions. Before the final version was completed, the whole instrument was checked by a Turkish language expert for clarity of language.

Data analysis

The quantitative responses were analyzed by using descriptive and inferential statistics. For these, the data were coded and prepared for analysis using the statistical analysis software SPSS 12.0. Both the means and the standard deviations of the questionnaire items were calculated. The inferential analysis was used to investigate the significant differences among mean scores of “teacher educators” and “deans.” For this reason, one-sample *t* tests were used.

The qualitative responses were analyzed by means of content analysis. Data reduction, data display, and conclusion drawing/verification phases were employed in this process (Miles & Huberman, 1994). This analysis began after the recorded interview sessions were transcribed into text. The interview participants first reviewed the accuracy of the details in the transcriptions of each interview. Data reduction activities included coding to represent, classify, and organize the data under the pre-identified categories and themes. Themes in the data were then identified, and the open-ended data from the questionnaires were coded and organized according to these themes. Data display refers to...
organizing and compressing information based on patterns and themes to permit conclusion drawing. During this phase, data based on themes was labeled and organized into data display matrices and structured summaries. Conclusion drawing and verification require the researcher to draw meaning from the displayed data. This final phase included noting comparing and contrasting, clustering, triangulation, and propositions. A synthesis of all the data gathered was reviewed and discussed by the researchers.

Results

The data were collected from: 1) deans, by means of five-point Likert-type scale questions and open-ended responses; 2) teacher educators, by means of five-point Likert-type scale questions, open-ended responses, and interviews; and 3) prospective teachers, by means of open-ended responses and interviews.

Analyses of the entire body of responses from all of the participants revealed a variety of main barriers and possible enablers. The researchers first identified and then categorized these responses. The results were grouped under the categories of main barriers and possible enablers, and then into two sub-themes under each category: quantitative and qualitative.

Main barriers

Quantitative results

The mean scores and standard deviations of the barriers perceived by the deans and the teacher educators are presented in Table 3. The findings indicate that a majority of the deans and teacher educators believe that “lack of in-service training,” “lack of appropriate software and materials,” and “lack of hardware” are all main barriers for integrating ICTs in pre-service teacher education programs. On the other hand, the following items fell below the mean, and a majority of the deans and teacher educators were neutral in their responses to the following items: “lack of appropriate course content and instructional programs,” “lack of time,” and “lack of appropriate administrative support.”

Table 3 also indicates the results of differences between two groups. One-sample t tests were conducted on the barrier items to evaluate whether the mean scores of the teacher educators were significantly different from the mean scores of the deans. From the statistical analysis, only two items significantly differed (p < 0.05): “lack of appropriate administrative support” and the overall barrier score, where the mean scores of teacher educators (2.94 and 3.79) were higher than the mean scores of deans (2.53 and 3.46). Although, these are significantly different scores, they must be interpreted carefully, since the number of teacher educators is double the number of deans.

<table>
<thead>
<tr>
<th>Barriers</th>
<th>Deans</th>
<th>Teacher educators</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Lack of in-service training</td>
<td>50</td>
<td>4.08</td>
<td>0.75</td>
</tr>
<tr>
<td>Lack of appropriate software/materials</td>
<td>49</td>
<td>3.81</td>
<td>0.93</td>
</tr>
<tr>
<td>Lack of basic knowledge/skills for ICTs</td>
<td>49</td>
<td>3.79</td>
<td>1.06</td>
</tr>
<tr>
<td>Lack of hardware</td>
<td>50</td>
<td>3.72</td>
<td>1.24</td>
</tr>
<tr>
<td>Lack of basic knowledge/skills for ICT integration</td>
<td>49</td>
<td>3.67</td>
<td>1.12</td>
</tr>
<tr>
<td>Lack of technical support</td>
<td>50</td>
<td>3.60</td>
<td>1.03</td>
</tr>
<tr>
<td>Lack of appropriate course content and instructional programs</td>
<td>50</td>
<td>3.38</td>
<td>0.97</td>
</tr>
<tr>
<td>Lack of time</td>
<td>50</td>
<td>2.60</td>
<td>1.16</td>
</tr>
<tr>
<td>Lack of appropriate administrative support</td>
<td>49</td>
<td>2.53</td>
<td>1.10</td>
</tr>
<tr>
<td>Overall</td>
<td>3.46</td>
<td>3.79</td>
<td></td>
</tr>
</tbody>
</table>
Qualitative results

According to the open-ended responses and interviews, and in agreement with the questionnaire responses, almost all of the participants believed that the items which were identified in the questionnaires, excepting “lack of appropriate administrative support,” were main barriers. In addition to the aforementioned barriers, the qualitative findings also revealed the followings as main barriers:

- crowded classrooms
- inadequate number of ICT-related courses
- lack of computers and other presentation equipment in classrooms
- lack of computer laboratories for use in free time
- lack of technology plans
- lack of motivation of the teacher educators concerning the use of ICTs in their classes
- lack of motivation of the prospective teachers concerning the use of ICTs in their courses and their future classes
- lack of good role models for prospective teachers
- lack of successful institutional models for STE.

One of the deans stated the following, concerning all of the barriers:

Teacher educators do not integrate ICTs into their classrooms due to disinterestedness which is caused by their insufficient ability and knowledge in the field. Another problem is overwhelming course load on teacher educators (that causes less time for research and personal development). Lack of time makes teacher educators stay away from ICTs, and they cannot fulfill their personal developments in this field.

One of the prospective teachers (A) said:

One month ago, I prepared my homework in CD format, but I could not show it to my teacher in class, because we don’t have any computers in the classroom. I want to present my homework through a computer by using flash animations and some pictures. Unfortunately, we have to prepare it by traditional methods. I think at least one computer should be placed in each classroom.

One of the prospective teachers (B) additionally noted a general need for computers, out of class:

We do not have any computer laboratories to use after the lesson. Sometimes, I have to go home to check my e-mail. Every time, there is lesson in the computer laboratory or it is closed. We could not use it after the lessons.

B also said the following in regard to teacher educators’ attitudes: “they [teacher educators] don’t have any positive attitude towards computers. If they had, they might be able to learn it. They cannot become a good model for the use of technology.”

On the other hand, the most important problem for one of the teacher educators was students’ attitudes. The teacher educator (C) indicated that:

We need to change the attitude of students in order to benefit from ICTs-related courses. They have negative attitudes and complain: “What will I do with it?” “Where will I use it?” “Why will I use it?” We should change these attitudes. This is the most important problem for me.

Possible enablers

Quantitative results

The means and standard deviations of possible enablers reported by the deans and teacher educators are provided in Table 4. There is strong agreement among the stakeholders; overall, they ranked “having technology plans” highest. The other leading enablers are “allocating more budget,” “allocating specific units and personnel for peer support,” and “offering in-service training.” On the other hand, “decreasing course load of the teacher educators” and “designing appropriate course content and instructional programs” items received the lowest mean scores as enablers Table 4 also indicates the results of the differences between two groups. One-sample t tests were conducted...
on the enablers’ items to evaluate whether the mean scores of “teacher educators” are significantly different from the mean scores of “deans.” Based on our analysis, none of the items significantly differed on mean scores ($p = 0.05$).

<table>
<thead>
<tr>
<th>Enablers</th>
<th>Deans</th>
<th>Teacher educators</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Having technology plans</td>
<td>50</td>
<td>105</td>
<td>0.324</td>
</tr>
<tr>
<td>Offering in-service training</td>
<td>50</td>
<td>109</td>
<td>0.189</td>
</tr>
<tr>
<td>Allocating more budget</td>
<td>50</td>
<td>105</td>
<td>0.263</td>
</tr>
<tr>
<td>Allocating specific units and personnel for peer support</td>
<td>50</td>
<td>105</td>
<td>0.095</td>
</tr>
<tr>
<td>Supporting teacher educators (i.e., incentive payment)</td>
<td>50</td>
<td>105</td>
<td>0.064</td>
</tr>
<tr>
<td>Decreasing the course load of the teacher educators</td>
<td>50</td>
<td>105</td>
<td>0.864</td>
</tr>
<tr>
<td>Designing appropriate course content and instructional programs</td>
<td>49</td>
<td>104</td>
<td>0.853</td>
</tr>
<tr>
<td>Overall</td>
<td>4.25</td>
<td>4.39</td>
<td>0.132</td>
</tr>
</tbody>
</table>

(levels of significance: $p < 0.05$ level: *)

Qualitative results

There is also a strong agreement between the open-ended responses and the participants’ interview results concerning possible enablers. A majority of the participants agreed on the items in the questionnaires, with the exception of “supporting teacher educators.” In addition to the aforementioned enablers, the qualitative findings revealed followings as possible enablers:

- having at least one computer in every classroom
- having at least one free laboratory in every STE
- supporting courses with an appropriate web page
- offering more ICT-related courses
- enhancing the motivation of the teacher educators and prospective teachers in regard to using ICTs in their classes
- designing ICT-related courses based on applicable activities
- being role models, as teacher educators, for prospective teachers by demonstrating how to use ICTs effectively in teaching.

One of the deans said the following about the enablers:

More teacher educators and technical support personnel need to be recruited at an adequate level. The current teaching staff must take in-service training under experts in the field. More competent STE should lead the less competent ones and transfer their experiences.

One prospective teacher said the following concerning the enablers: “Instructors should be provided with in-service training for the integration of ICTs.” On the other hand, another interviewee suggested a proficiency exam on ICTs for the teacher educators. If a teacher educator takes this exam, then s/he should be motivated (i.e., by financial incentives).

In addition to the themes stated in the questionnaire, one teacher educator suggested that the prospective teachers’ motivation should be enhanced. He offered further suggestions for new ICT-related courses for the STE. According to him, two ICT-related courses were not meeting the needs. Also one of the teacher educators (C) commented:

ICT-related courses should be integrated into school-experience courses. I think this model would enhance the efficiency of the integration. Using ICTs is important, but the integration of ICTs in your class is more important. Also, we have to offer our students a new ICT-related course which has to include both ICTs and a field of study (e.g., math, language, chemistry) after the method courses.
Discussion

Generally, there was agreement among the results of deans, teacher educators, and prospective teachers’ concerning the main barriers and possible enablers to successful ICT integration in pre-service teacher education programs. According to Willis (2001), enablers are local, not universal; however, the findings of this study show that not only the enablers but also the barriers are similar to those identified in the literature, as summarized in Table 1 and Table 2. This study also indicates contradictory results compared to the literature on barriers (USDE, 2000) in that “lack of appropriate course content and instructional programs,” “lack of time,” and “lack of appropriate administrative support” were below the mean, and a majority of the participants identified these statements as not representing main barriers.

It is important to recognize that a number of factors have been identified which encourage and enable all of the stakeholders to integrate ICTs in pre-service teacher education programs. The findings revealed that the following strategies could provide a generic approach towards enhancing this ICT integration: technology plans, in-service training, strong infrastructures, technical support, and role models.

Technology plans

The questionnaire results of the deans and teacher educators showed that a technology plan could be one of the possible enablers. The strategy “having technology plans” ranked as the most commonly agreed-upon item by both groups. The effective integration of ICTs in STE is possible if future goals and strategies are set and implemented in a planned manner. According to Bates (2000), ICT infrastructure needs to be guided by the administrative, financial, and teaching needs of the STE, which in turn are reflected in the technology plan that should integrate the vision and strategic direction of the institution. In this sense, the first possible enabler is to develop a technology plan for the STE. Every STE can prepare a technology plan, and they can employ a technology support task force for both technical and instructional purposes. Existing plans, policies, and strategies additionally need to be updated, developed, and spread to all stakeholders. This idea was also supported by Anderson, Varnhagen, and Campbell (1998); Fabry and Higgs (1997); Moursund and Bielefeldt (1999); Rogers (2000); and UNESCO (2002).

In-service training

According to the participants, teacher educators need leadership and require training in methods for integrating ICTs in their classrooms. Thus, the present research results are parallel to the literature (Rogers, 2000) and lend support to the strategy of appropriate in-service training for teacher educators. Almost all of the participants noted “lack of in-service training” as a main barrier.

Instructional technology centers can be founded in universities to lead the departments to use ICT tools effectively and integrate them into an educational environment, as well as to offer in-service training. Moreover, these centers can organize and decide which ICT resources will be purchased and how available resources could be used most efficiently. These centers can also serve to offer peer support to the teacher educators, and can further offer to the public the use of existing ICT resources.

Strong infrastructures

The findings of previous research (Anderson, Varnhagen, & Campbell, 1998; Cuban, 2001; Cuban, Kirkpatrick, & Peck, 2001; Ertmer, 1999; Schoep, 2004; Vaughan, 2002) and the findings of this study suggest that providing access to ICTs is not enough. ICT resources and infrastructures in STE are limited. STE should invest in larger budgets for purchasing new hardware and software, and for updating and upgrading them. Therefore, funding for new ICT resources should be increased in order to provide adequate ICT equipment and resources.

Rather than limiting ICTs to certain centers (laboratories) and courses (ICT-related courses), ICTs can be spread to the whole physical environment of an STE such as canteens, corridors, and particularly classrooms and courses from the introductory to the school-experience level. This would create a more authentic environment and involve students.
in more practice. Moreover, laboratories can be kept open for the use of students not only during lesson hours but also after lessons by employing student assistants.

With the aim of increasing quantity of instructional software and materials, cooperation between the STE and related companies can be useful in providing software. For instance, portals that function as reusable learning objects can be formed. Projects or objects can be developed and uploaded to those portals to provide a discussion environment (Klinger, 2006). In this process, graduate students of Instructional Technology departments may provide necessary support, as they can work as professional instructional technologists. K–12 teachers can be participants in these portals as well. This can enhance the cooperation among all participating institutions to better ensure successful ICTs integration.

Technical support

In this study, specific units and personnel were advocated for technical support to teaching staff as a cost-effective method. Universities and/or STE can allocate specific units or personnel to provide technical support, to assist with the public’s use of ICT tools and materials in instruction, and to help reduce the instructor workload. The aforementioned instructional technology centers can also offer technical support. This finding was supported by Ronnvist, Dexter, and Anderson (2000); Rogers (2000); and Sandholtz (2004).

For ICT issues, necessary policies can be constructed by cooperating with other institutions (e.g., employers, universities, STE) in order to supply personnel such as technicians. They can provide timely training and arrange peer collaboration. Peer support and technical support might be chosen as in-service training methods.

Role models

The findings in the literature parallel this study’s results in terms of the advocated need for good role models. Prospective teachers should observe appropriate models throughout their undergraduate process (Bullock, 2004; Hornung & Bronack, 2000; Kariuki, Franklin, & Duran, 2001; O’Bannon, Matthew, & Thomas, 1998; SITE, 2002; Yildirim, 2000; Whetstone & Carr-Chellman, 2001).

It is important that teacher educators act as role models for prospective teachers by using ICTs. They can demonstrate their competency and willingness to use ICTs in teaching and should use ICTs in their classrooms. Other than basic ICT applications (MS Office), they need to be aware of other appropriate software (e.g., tutorials and simulations) and use these programs to enrich their courses in an ICT-integration process.

Conclusion

Even though the sample in this study was limited to 18 STE for teacher educators and prospective teachers and convenience sampling with representative methodologies was used for both groups, this study does provide a good picture of the views of deans, teacher educators, and prospective teachers pertaining to the main barriers and possible enablers for ICT integration in their schools. To create an environment of effective ICT integration, pre-service teacher education programs must focus on eliminating barriers. Based on the findings and discussions presented here, the several recommendations are offered for practitioners. Future research is needed to verify the effectiveness of the following recommendations and to identify other important ones:

1. Technology plans for implementing ICTs in STE should be prepared and implemented.
2. Specific units and personnel should be allocated for peer support and organization, as well as to assist in the public’s use of ICT tools and materials for ICT-enhanced instruction.
3. The teacher educators who integrate ICTs in their courses should be supported (i.e., through incentive payments).
4. The course load of teacher educators should be decreased.
5. Teacher educators should act as role models for prospective teachers by using ICTs in their courses.
6. In-service training in ICTs for teacher educators should be improved in both quantity and quality.
7. Every classroom should have at least one computer with Internet access and an LCD projector.
8. Every STE should have at least one laboratory available to students.
9. Course content should be redesigned to acquire more benefit from ICTs.
10. Courses could be supported by a course delivery system (e.g., LMS, course support web page).
11. More ICT-related courses for prospective teachers should be offered.
12. Every ICT-related course should be based on practice-oriented.
13. ICT-related courses should be integrated in teaching practice courses.
14. A new ICT-related course, which must include both ICTs and a field of study (e.g., math, language, chemistry), should be integrated in the curriculum after the method courses.
15. Teacher educators and prospective teachers should be aware of the benefits of ICTs.

References


Utilizing Computer-mediated Communication Tools for Problem-based Learning

Hao-Chang Lo
General Education Center, Chung Shan Medical University, Taiwan // Tel: +886-4-24730022-17123 // Fax: +886-4-23248184 // simonlo@csmu.edu.tw

ABSTRACT
This study aims to strategically use computer-mediated communication (CMC) tools to build online communication environments for problem-based learning (PBL). A six-stage process was proposed for online PBL learning in this study: 1) identifying the problem, 2) brainstorming, 3) collecting and analyzing information, 4) synthesizing information, 5) co-building knowledge, and 6) refining the outcomes. The one hundred undergraduate students who participated in this study were randomly grouped into 13 groups. Multiple methods of data collection were adopted to investigate students' learning activities in the online PBL course. The methods of data triangulation and investigator triangulation were used to analyze the following: 1) the use of CMC tools, 2) students' learning achievement, 3) students' experience in the online PBL course, and 4) group learning activities. The empirical evidence showed that students were able to communicate, discuss, and co-build the knowledge from the collected information. With the online communication, they were able to seek solutions to the problems in learning activities. Meanwhile, students were satisfied with the online PBL course. The successful experience of course design in this study can encourage instructors to use open-source or free-hosted CMC services to develop online PBL courses.

Keywords
Blog, Computer-mediated communication (CMC), Microsoft Network (MSN), Problem-based learning (PBL), Wiki

Introduction
Problem-based learning (PBL) was derived from the research project conducted at the Department of Medicine, McMaster University (Canada). PBL is an approach emphasizing students' learning through active inquiry in small groups (Sonmez & Lee, 2003). In a review of research, Schmidt and Moust (2000) concluded that PBL follows the "seven jump" procedure in a typical tutorial group, in which students do the following: 1) clarify unknown terms and concepts in the problem description; 2) define the problems; 3) analyze the problems by brainstorming, in an attempt to produce as many different explanations as possible based on prior knowledge and common sense; 4) criticize the explanations proposed and try to produce a coherent description of the processes; 5) formulate learning issues for self-directed learning; 6) fill the gaps in their knowledge through self-study; 7) check whether they have learned, share findings with peers, and try to integrate the knowledge acquired into a comprehensive explanation for the problem.

Many studies provide evidence of the effects of PBL in terms of improving students’ learning, and PBL has been adopted widely in educational settings (Savery & Duffy, 1995). PBL can foster students’ intrinsic motivations and develop their self-learning skills because it provides a student-centered and self-oriented learning for students to seek solutions to real world problems (Barrows, 1997; Gallagher, Sher, Step, & Workman, 1995; Prince & Felder, 2006). Norman and Schmidt (1992) indicated that the learning in a PBL format can lead to a long-term retention of knowledge, enhance the integration of basic science concepts into clinical problems, and result in an increase in intrinsic interest in the subject. Antepohl and Herzig (1999) suggested that students consider PBL to be an effective learning method and favor it over the lecture format. Furthermore, students reported positive effects of PBL in terms of use of additional learning resources, interdisciplinary courses, team work and learning fun. Tiwari, Lai, So, and Yuen (2006) proposed that those who undertook PBL are significantly better in the development of critical thinking dispositions, as measured by the California Critical Thinking Disposition Inventory (CCTDI), than those who took lecture courses.

With the advancement of technology, small-group discussion in the classroom can be replaced by transmitting messages via networked computers. Such computer-mediated communication (CMC) allows interactions among geographically separated students, who can communicate and learn through dialogue exchanged on the Internet. Online courses are, therefore, designed with learning activities for small groups of five or six heterogeneous
members who collaborate or cooperate to complete a common task (Henri & Rigault, 1996). The majority of these online courses use constructivist educational principles (Bangert, 2004). Using CMC in the PBL process was proposed and defined as an approach to distributed problem-based learning (dPBL) (Cameron, Barrows, & Crooks, 1999).

Both teaching and learning with the dPBL approach have greatly contributed to the inquiry into the use of CMC (e.g., ChanLin & Chan, 2007; Steinkuehler, Derry, Woods, & Hmelo-Silver, 2002; Oberlander & Talbert-Johnson, 2004). There are various forms of CMC, such as instant messaging, peer-to-peer networks, e-mail, bulletin boards, online chat rooms, and massive multiplayer online (MMO). Usage is either synchronous or asynchronous. CMC can be accessed by commercially available applications like Blackboard and WebCT, which are designed particularly for teaching. Open sources such as MSN, blogs, and wikis can be designed to manage online learning. These CMC tools can be user-friendly communicative environments for today’s students, who are familiar with the use of CMC. Based on accessibility, feasibility, and cost-effectiveness, this study aims at strategically using open-source software as CMC tools to build an online course for PBL.

Methods

Participants

One hundred undergraduate students from nine departments (66 freshmen, 10 sophomores, 17 juniors, and 7 seniors) chose the two-credit, 15-week, web-based elective course, a general education course called Applied Science in Life. They were randomly divided into 13 groups with an average of six or seven members per group. The instructional objective of this course was to develop students’ abilities in constructing new knowledge through the researcher-designed PBL approach.

CMC environment for the online PBL classroom

Figure 1 shows the design of a CMC environment for the online PBL classroom in this study. The online PBL classroom needs CMC tools for the following: 1) synchronous discussion in the process of problem solving; 2) asynchronous communication that allows students to take their time to read, write, and give comments to members within a group; and 3) a platform that allows students to collaboratively edit their group project and to learn from other groups. Therefore, three tools were used: instant messaging (IM) for synchronous discussion, blogs for intra-group communication, and wikis for collaborative group work and inter-group learning.

MSN (or MSN Messenger) is one of the most popular IM interfaces for college students in their daily lives. In this study, students invited the group members to join their learning community so that group members could communicate synchronously in the PBL process if necessary. Students were requested to invite the research-designed virtual contact, called the PBL recorder, to record what they had discussed.

A blog is a website that contains frequently updated posts, with the most recent entry at the top of the page and the previous ones displayed in reverse chronological order. This study set up one blog for each group (Figure 2) using
Movable Type Open Source 4.1 (http://www.movabletype.org/). Group members could asynchronously post the information they find, make comments on learning issues, incorporate hyperlinks to other blogs or news sources, and discuss related topics. This means that a blog was the asynchronous intra-group communication tool while members were seeking data to solve a problem.

A wiki is software that allows users to easily create and edit collaboratively. This study used open source MediaWiki (http://www.mediawiki.org/wiki/MediaWiki), free server-based software whose site gets millions of hits per day. Group members collaboratively wrote group reports on the wiki (Figure 3) after they had discussed each learning issue on the blog. Further, the group wiki reports could also be reviewed by other groups. That is, wiki became the collaborative documenting tool for group members and the group presentation interface for each group to review.
Course design

In the first two weeks, students came to the conventional classroom for the orientation on how to work online in a group. The web-based course began in the third week to learn six issues in 12 weeks, that is, two weeks for each learning issue. The learning issues include topics of housing, food, clothing, household goods, transportation, and entertainment. In the fifteenth week, the teacher and students came to the classroom for face-to-face feedback, evaluation, and final discussion.

For each learning issue, a six-stage process was proposed as follows: 1) identifying the problem, 2) brainstorming, 3) collecting and analyzing information, 4) synthesizing information, 5) co-building knowledge, and 6) refining the outcomes:

- At the first stage, the problem for the learning issue was posted on wiki in the form of text, pictures, sound, video clips, or hyperlinks for students to access the original information. The learning issues, selected from newspapers, magazines, TV news, or Internet news, were current issues that do not have just one answer or one definite solution. For example, one of the learning issues is that a renewable fuel, biodiesel, has been produced to reduce serious air pollutants, and therefore, is expected to reduce global warming. Students were asked to answer whether they would choose biodiesel if they had a car.

- At the second stage, students needed to brainstorm or generate ideas. They could use MSN to discuss online in real time with fellow group members. They needed to tackle the problem through self-directed questioning. For example, in order to answer whether they would use biodiesel, students needed to raise questions like the following: What is biodiesel? How does the diesel engine work? What is the difference between biodiesel and conventional fuel? Is biodiesel good for the diesel engine? How much should we pay for biodiesel? This kind of self-regulated questioning was aimed to arouse students’ intrinsic motivation.

- At the third stage, group members were assigned to collect information. They had to fulfill their own task by the deadline that was agreed upon. They discussed via MSN or posted on the blog what they found and what they learned. Thus, they collaboratively collected useful information.

- At the fourth stage, the collected information was critically analyzed so that the relevant data could be synthesized to solve the problem. Generally speaking, knowledge building occurred in the process of collecting, analyzing, and synthesizing stage.

- At the fifth stage, students presented the solution to the learning problem/issue collaboratively on wiki. Students reflected and elaborated on what they had learned.

- At the final stage, the course instructor gave feedback and suggestions to help students improve. For example, they might be directed to another group report on wiki. In this way, students learn from other groups and are not limited in their own group.

To assist the implementation of learning activities, especially in the negotiation of the time of online discussion, group members elected a leader. They established networks of communication in their group by e-mail, cell phone, and MSN. The group leaders had to stay connected throughout the networks of communication to ensure that group members were present at the learning activities. The group leaders would inform the tutor if the absences were frequent.

The course instructor and a research assistant played the roles of tutors in the PBL work. However, they did not play the role of experts as tutors in the conventional PBL process did. Instead, they guided students as supervisors and facilitators in the learning activities. The research assistant monitored the learning process for nine hours a day (i.e., three hours in the morning, afternoon, and evening). At the final stage, the course instructor concluded the learning issues with feedback and suggestions but no answers.

Multiple assessment methods were adopted to evaluate students’ learning, including 1) the achievement test (20%), 2) the group report (20%), 3) the self and peer assessment (30%), and 4) the interactions intra-group (30%). The achievement test was based on knowledge of the learning issues. The group report was graded by the course instructor according to the content and the inferences that the group members documented on the wiki. The self and peer assessment was a 5-point Likert scale (Cronbach α = 0.9198) with 11 items (four items for self evaluation and seven items for peer evaluation) developed by Chung Shan Medical University in 2005 (Chen, Lee, Lee, Wang, Lin, & Yang, 2006). The course instructor who observed their performances in the process of discussion for each learning issue evaluated students’ interaction in the group, scoring them from 0 to 5.
Data collection and analysis

Multiple methods of data collection were adopted to investigate the students’ learning activities in the online PBL course via computer-mediated communication tools and to ensure that interpretations of the data will be both reliable and valid in this study. According to the triangulation proposed by Denzin (1978), investigator triangulation and data triangulation (as shown in Figure 4) were designed in this study. Multiple sources of data, as shown in Table 1, were collected, and data triangulation was used to analyze these multiple sources of data. Investigator triangulation was used to synthesize the observations of the course instructor and the research assistant. As shown in Figure 4, two tutors acting as investigators analyzed the multiple sources data to explore clues to understanding students’ learning behavior in the online PBL course. They interpreted the raw data separately first and then presented the interpretations. To clarify the dubious data and different interpretations, they had to interview the students.

Table 1. Multiple sources of data collection

<table>
<thead>
<tr>
<th>Sources of data</th>
<th>Means to collect required data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialogue in virtual classroom</td>
<td>The system recorded all data on the blog, wiki, or MSN for the duration of the learning process and they will be encoded according to the sequence of time.</td>
</tr>
<tr>
<td>Questionnaire (Q)</td>
<td>Students completed a questionnaire using a 5-point Likert scale at the end of the course. The aim of the questionnaire was to examine students’ experience participating in course activities and using learning interfaces.</td>
</tr>
<tr>
<td>Interviews (I)</td>
<td>Tutors used interviews when they noticed something unusual in the learning activities or when different interpretations occurred in data analysis.</td>
</tr>
<tr>
<td>Field notes (Fn)</td>
<td>The tutors, as observers and participants, recorded meaningful events for the duration of the study. Field notes include date, time, specific facts, sensory impressions, and immediate reflections or thoughts.</td>
</tr>
<tr>
<td>Reflection on learning experience (RLE)</td>
<td>Students and tutors met in a traditional classroom and shared their learning experience with one another at the end of the course.</td>
</tr>
<tr>
<td>Assessment of students’ learning (ASL)</td>
<td>Multiple assessments consist of the achievement test, the group report, the self and peer assessment, and the interactions intra-group.</td>
</tr>
</tbody>
</table>

Figure 4. Illustration of data triangulation and investigator triangulation
Results

The use of CMC tools

Based on the questionnaire responses, the usage frequency of the three CMC tools shows that MSN was the most popular tool. Most students (96%) were familiar with MSN, and almost all of them (95%) did not have any difficulty using MSN in the process of learning. More than half of the students (63%) were already familiar with blogs prior to the course, but only a few students (10%) knew about wiki. After receiving training in the use of CMC tools in the first two weeks of this course, students showed significant growth in confidence and familiarity with using blogs (81%) and wikis (58%). This indicates that the training in CMC tools was necessary for students to ensure that they were able to participate in the learning activities via these communication tools.

When comparing the synchronous MSN tool with the asynchronous blog for the intra-group discussion, data from field notes (Fn) and reflection on learning experience (RLE) revealed that students tended to discuss in real time. Subsequent to the offline discussing, they did not post the unfinished discussion on the blog. Instead, they preferred to continue the instant discussion. Students said the following:

- “Discussion was often unfinished, when group members were offline and did not continue post it to the blog …” (Fn 950422).
- “I prefer to continue the discussion when they are online next…” (RLE 941803401).

It is likely that students prefer to communicate in the environment that they are familiar with. However, the discourse via MSN was more like what they would say in a conversation, rather than what they would reflect in a piece of writing. For example:


In contrast, data of the asynchronous CMC tools both on the blog and wiki shows that students worked on content and built more organized discussions:

- “Biodiesel is an alternative fuel which derived from the oils and fats of plants and is a valuable form of renewable energy. Biodiesel can be used directly in any existing, unmodified diesel engine like conventional diesel fuel. That is, it will help reduce the countries’ dependence on foreign oil imports if we can produce biodiesel in Taiwan” (DVC G3T1842119).

Students’ learning achievement

The results of achievement tests showed that the difference of the pre-test ($M = 31.70$, $SD = 9.95$) and post-test ($M = 61.55$, $SD = 9.12$) is significant ($t = 26.127$, $p < .001$). This indicates a significant gain in knowledge relevant to the learning issues after the PBL online course. Meanwhile, the results of the self-assessment and questionnaire responses showed high percentages (76% and 86% respectively), and agreed that they acquired knowledge of learning issues in the process of online PBL course.

The group reports presented on the wiki provided evidence that students could solve the problems via the process of the online PBL course. For instance, all groups concluded whether or not they would use the renewable fuel based on the analysis of biodiesel:

- “Biodiesel is a form of diesel fuel manufactured from vegetable oils, animal fats, or recycled restaurant greases. It is safe, biodegradable, and produces less air pollutants than petroleum-based diesel … To sum up, it is an innovation energy which fits in with economic benefits and is good for the environment. We think that we will use the renewable fuel if possible” (DVC G3T2041209).
- “Though biodiesel has a lot of advantages, it has its limits. First, some vehicle manufacturers worry about the use of biodiesel and remain cautious over the use of biodiesel. In the UK, for instance, many only maintain their engine warranties for use with maximum 5% biodiesel (blended in with 95% conventional diesel). Secondly, diesel engine cars are more expensive than conventional ones (so far in Taiwan). It means that we have to pay more money for the car but the effects on environmental protection and economy are limited. Hence, we do not think that we will choose it as fuel of car unless the expense is subsidized by the government” (DVC G9T5041211).
Students’ experience with the online PBL course

Most of the students (76%) thought that it was more interesting to learn in a virtual classroom than in a face-to-face classroom. Students could synchronously search and renew information during the discussion. They were satisfied with the communication interfaces that could provide permanent links to web pages. It was likely that they liked learning in a virtual environment because of the flexibility and the functions of CMC, as demonstrated by the following comments:

- “There was something different … it did not seem to be a formal course because the experience was so different from that of learning in a traditional classroom” (RLE 947702402).
- “It was good that I could search necessary information when I discussed with group members” (RLE 947701705).
- “It was convenient that we could read the contents of online information via hyperlink posted on MSN or blog whenever we needed” (RLE 946023301).

However, some students (12 out of 100) felt a sense of unreality and distance in the virtual classroom. As one student stated:

- “The atmosphere of discussing the issues in a learning community in a computer-supported environment was something unreal … and it distanced us … Although discussing in a face-to-face classroom was more serious and formal…” (RLE 945204001)

Group learning activities

The results of the responses of the peer assessment showed a high percentage of satisfaction with group learning. Most students agreed that group participation was good (81%); peer interaction was good (84%); the content of the discussion was systematic, well-organized, and substantial (80%); the group members all strived to search and collect information (75%); group members were keen on PBL (81%); and group members acquired knowledge to meet learning objectives (86%). According to the attendances, dialogues, and interactions in a group, students had scored with a range from 0 to 10 for each item. The researcher was satisfied with students’ performance in the process of learning because students had scored an average of 22 (with a range from 0 to 30).

Most of the group leaders (9 out of 13, or about 70%) successfully maintained the flow of the learning activities throughout this study. Moreover, a competent leader could enhance learning activities, as one student commented on his group leader:

- “I thought that we did a good job. Thanks to our group leader … We just needed to work according to the plan made by the group leader” (RLE 942423203).

Unfortunately, learning activities in three groups were chaotic for lack of leadership. One student sounded anxious about his absent leader:

- “Where is the leader? Can someone get in touch with him? We cannot present the report in time. What shall we do?” (DVC G4T2030501).

Even though the absence of leadership might cause difficulties for group learning, one group did manage to work without a leader. One member from the group seemed to be calm and encouraged others to continue the group work:

- “Ok! Let’s discuss it ourselves. We are going to assign learning tasks if the leader is absent after two days. What do you think?” (DVC G7T1231104).

Generally speaking, the leader plays an important part in an online learning group. An incompetent leader always causes difficulties for group learning unless an alternate leader emerges.

Conclusions and suggestions

In this study, the researcher developed an online PBL course which was supported by CMC tools of open-source software. MSN, blogs, and wikis were strategically used to build an online environment for students to learn via the PBL approach. With the training in MSN, blogs and wikis they had received prior to participating in web-based
learning, students were able to communicate, discuss, and co-build knowledge from the collected information. Through online communication, they were able to seek solutions to problems in learning activities. Moreover, they perceived participation and interaction in a group as motivation to learn. In general, students were satisfied with the online PBL course. The cost-free CMC tools have proved able to serve the purpose of supporting PBL. The successful experiment of course design in this study can encourage instructors to use open-source CMC or free-hosted CMC services to develop online PBL courses. Effectively integrating these separate CMC tools into a unique interface will be the next task.

This online PBL course is designed to be student-centered for self-oriented learning. To further improve the design of web-based learning, the findings of this study suggested taking blended-learning into consideration because the sense of unreality and distance exists in the virtual classroom. A blended approach with appropriate face-to-face encounters is deemed to be more relevant for students’ needs in terms of student-centered learning (Donnelly, 2004). In this study, students presented and shared their group reports in a face-to-face classroom environment at the final stage of PBL. Rovai and Jordan (2007) provided evidence that blended courses would produce a stronger sense of community among students than a traditional or a fully online course do. For further design of the online PBL course, variations in disciplines need to be considered. For example, when particular professional knowledge, such as the medical field, is crucial in the PBL courses, tutors must have inter-disciplinary expertise.

Acknowledgements

This work is supported by the National Science Council, Taiwan, Republic of China, under contract NSC94-2516-S-040-003.

References


Examining Applicants for Admission and Completion of an Online Teacher Certification Program


Texas A&M University, College Station, Texas 77843-4232 USA // jdenton@tamu.edu // trinadavis@tamu.edu // rcapraro@coe.tamu.edu // ben_smith@tamu.edu // lbeason@coe.tamu.edu // dgraham@tamu.edu // strader@tamu.edu

*Corresponding author

ABSTRACT

The purpose of this research was to determine whether particular biographic and academic characteristics would predict whether an applicant would matriculate into and successfully complete an online secondary teacher certification program. Results of statistical analyses conducted on 170 applicants over the past three years affirm the importance of the applicant’s performance on the state content licensure examination. Other academic variables and biographic variables were found not to influence success in completing this secondary teacher certification program.

Keywords

Online Teacher Certification, Teacher Preparation, Applicant Profiles, Alternative Certification, Pathways to Teaching, Alternate Certification Routes

In the introduction of their monograph, A Good Teacher in Every Classroom, Darling-Hammond and Baratz-Snowden (2005) note that more than one hundred thousand new teachers enter classrooms across the country each year. These beginning teachers vary greatly in the professional skills and knowledge they bring to the classroom with many having little or no formal learning experiences regarding essential knowledge and clinical experiences needed to succeed in the classroom. Darling-Hammond and Baratz-Snowden continue by listing the following traits exhibited by teachers whose students exhibit strong academic achievement.

Effective teachers use many different tools to assess how their students learn as well as what the students know. They use this information to help students advance from where they are to where they need to be. They carefully organize activities, materials and instruction based on students’ prior knowledge and level of development so that all students can be successful. They know what conceptions students should bring with them about the subject and what misconceptions are likely to cause them confusion - and they design their lessons to overcome these misinterpretations. They adapt the curriculum to different students’ needs; for example, making content more accessible for students who are still learning English and for those who have special educational needs. (pp. 2 -3)

These authors stress that teacher preparation has evolved to the degree that there is agreement on what beginning teachers need to know in order to enter the classroom with professional competence to adequately serve the very first students they teach. Two other recent policy papers supported by the National Council on Teacher Quality and the Education Commission of the States provide accounts regarding whether empirical research supports commonly held notions and practices applied to teacher certification, selection, retention and compensation. The reports, Increasing the Odds: How Good Policies Can Yield Better Teachers (National Council on Teacher Quality, n. d.) and Eight Questions on Teacher Preparation: What Does the Research Say? (Allen, 2003) examine factors affecting policies enacted by state governments and local school districts in deciding who will be admitted to the teaching profession. Drawing from these sources, it is evident the assumed relations among variables (professional and personal characteristics) for selecting and preparing tomorrow’s teachers need additional examination given the following findings reported in these reports.

- Research evidence does not exist to support or refute more stringent teacher preparation entrance requirements or conducting more-selective screening of program candidates to positively affect student achievement.
- Personal attributes of high achievement orientation, accepting responsibility, demonstrating critical thinking, being organized, being motivated, being respectful of others, and supporting the goals of the organization are soft attributes thought to relate to greater student academic gains.
• Teachers with strong academic credentials generally produce greater student academic gains.
• Empirical evidence supports the idea of a positive association between the level of literacy of teachers and academic achievement of their students.
• Subject matter knowledge has mixed empirical support for secondary mathematics and science teachers with greater subject preparation generally yielding a positive relation to student achievement, but this modest support does not occur with other secondary content areas.
• The research evidence suggests that preparation in a given subject does not necessarily develop understanding of how particular knowledge and skills of the content domain are best learned.
• There is insufficient empirical evidence to support any teacher preparation models that will produce effective teachers in hard-to-staff and low-performing schools.

The underlying theme from all of these reports is the call to strengthen research capacity by defining a strategic and coordinated research agenda for teacher preparation. Our experiences (Denton, et al. 2004-05) with technology professional development for teacher educators and classroom teachers are consistent with the literature (Garet, Porter, Desimone, Birman, & Yoon, 2001; Joyce & Showers, 2002; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003) regarding professional development experiences that emphasize academic subject matter (content), provide opportunities for “hands-on” activities (active learning), are integrated with ongoing classroom operations (coherence) and provide many development experiences for an extended period of time. These experiences are more likely to produce desired knowledge and skill changes. In melding our experiences with an evolving research agenda, we begin by examining academic and biographic variables of applicants associated with being admitted as a candidate and then completing all requirements for secondary teacher certification. The following sections provide the research questions, a literature review of the continuing need for well prepared science and mathematics secondary teachers, a description of alternative certification programs, a description of components of the Accelerate Online/OPTIONS program including recruitment and placement protocols, and then proceeds to the methods, results and conclusions.

Research Questions

The following research questions have been phrased to guide this inquiry.
1. What academic and biographic variables of program applicants relate to successful admission into the certification program?
2. What academic and biographic variables of program candidates relate to successful program completion and certification?

Related Literature

Teacher Shortage or Turnover

Recurring accounts of local school districts and state departments of education experiencing teacher shortages continue to appear in professional journals (Budig, 2006; Guarino, Santibanez, & Daley, 2006), daily newspapers (Chaker, 2006; Houston Chronicle, 2007), and the evening news offered by broadcast networks. To illustrate these accounts, approximately half of our nation’s promising young educators, as well as their more experienced colleagues, choose to leave their schools during their first five years in the profession (Latham & Vogt, 2007; Levin, 2006; Weaver & O’Brien, 2004). Of those who stay, especially teachers in high need settings, some remain to gain the seniority needed to transfer to schools whose needs are not as pronounced. While not affecting the overall teacher supply in a state, transfer from high need settings can and do lead to continuing shortages of experienced teachers in high-turnover schools (Hull, 2004).

Teacher turnover can be grouped under two categories, migration and attrition, with about equal numbers of exiting teachers grouped under each category. Ingersoll (2003) reports that teachers usually offer personal reasons for exiting their teaching position, but some (about one-fourth of those leaving) add that job dissatisfaction due to low salaries, lack of instructional and emotional support, feelings of isolation, and little influence on how they do their work are reasons for their departure from classrooms. Approaches frequently noted in the literature for reducing
teacher attrition include quality induction and mentoring programs (O’Neill, 2004) while the teacher migration problem can be addressed by providing extended professional development opportunities for all teachers beginning with induction and continuing with career-long opportunities for instructional leadership throughout their careers (Hull, 2004).

Quality Teacher Preparation

Factors contributing to national leaders’ concerns about teacher supply are that states and school districts face the challenge of attracting more people into teaching while improving teacher quality. The report, Meeting the Highly Qualified Teachers Challenge: The Secretary's Annual Report on Teacher Quality (United States Department of Education [USDE], 2002), provides recommendations for carefully examining and changing aspects of teacher certification including: state certification requirements, minimal criterion teaching candidate performance on state licensure examinations and the practice of hiring teachers on emergency certificates. The following quotation captures the serious tenor of this report about teacher certification and teacher education programs across the country.

Many academically accomplished college graduates and mid-career professionals with strong subject matter backgrounds are often dissuaded from entering teaching because the entry requirements are so rigid. At the same time, too many individuals earn certification even though their own content knowledge is weak (p. vii).

Academic standards for teacher certification are set quite low. To illustrate, for a teacher licensure test, only 1 of 29 states sets its passing score near the national average in reading; 15 states established passing scores below the 25th percentile. Similar low criterion pass scores are noted for mathematics and writing tests. These low criterion scores and the statistic that more than 90 percent of teaching candidates pass these tests do not instill confidence in the quality control aspect of licensure examinations (Goldhaber & Brewer, 2000; USDE, 2002). Supporting this position, Darling-Hammond (2006) describes scenarios where individuals with no academic preparation in teaching have performed as well on standardized tests of teacher competency as trained teachers. This situation has occurred because of poorly conceived items that do not represent the professional knowledge base of pedagogy and poorly constructed items where “correct” answers can be determined by a careful reading of the items.

Higher Standards

The demand for higher science education standards is evident at the state and local levels as well. A few years ago Texas added science (biology, integrated chemistry, and physics) to the Texas Assessment of Knowledge and Skills (TAKS) examination that 11th-grade students are required to pass in order to graduate from high school (Office of Texas High School Education, 2001). Yet in order to meet this graduation requirement, a major challenge lies in the shortage of qualified science teachers (Chaika, 2005; Darling-Hammond, 2000a; National Center for Education Statistics, 2001; National Research Council [NRC], 2000a, 2000b; USDE, 2000; Urban Teacher Collaborative [UTC], 2000). “Research has generally shown that high school math and science teachers who have a major in the subjects they teach elicit greater gains from their students than out-of-field teachers…” (USDE, 2002). Supporting this point-of-view for more content, Goldhaber and Brewer (2000) reported from their analysis of a subset of the national Educational Longitudinal Study of 1988 that 12th grade student achievement in mathematics was greater if their teachers had subject specific preparation (a mathematics degree or teacher certification in mathematics). Most experts agree that personnel entering science teaching fields should have a science major in addition to pedagogical studies (American Federation of Teachers, 2000; Berry, 2001; Bybee & Loucks-Horsley, 2001; NRC, 2000a). Further, Darling-Hammond (2000b) emphasizes the importance of pedagogical studies in her analysis of teacher quality on student achievement citing research that reported consistent positive effects of teachers’ formal education preparation on teaching performance and student learning. Supporting this position, Cochran-Smith (2002) states that quality formal teacher preparation includes instruction on many processes including: representing complex knowledge in accessible ways, asking good questions, forming relationships with students, interpreting multiple data sources, posing and solving problems of practice.

Nevertheless, the use of non-credentialed personnel in teaching roles was widespread in urban school districts a few years ago (UTC, 2000), with 20-25% of high school science and mathematics teachers nationwide not possessing a
major or minor in their teaching field (USDE, 2000). Unfortunately, academically weak college graduates are more likely than those who are academically strong to prepare to teach and to remain in teaching (USDE, 2002).

**Advancing Knowledge and Understanding in Preparing Natural Science Teachers**

Prior reports have documented the need for fundamental restructuring of teacher preparation and professional development. In addition, a potential solution consists of directing the energy and talent of graduate students and postdoctoral scientists to reinvigorate science education in schools (Alberts, 1999). This requires creating new, more accessible pathways for science professionals to enter teaching careers. One study included the finding that while 36% of doctoral-level mathematics and science graduate student/postdoctoral fellows surveyed had considered secondary school teaching in their career decision-making, only 0.8% of science and mathematics Ph.D.s work in K-12 education. Reasons for not pursuing teaching careers include perceptions about the difficulty of obtaining teacher certification (NRC, 2000b). Paradoxically, many undergraduate teacher preparation programs produce teachers who are weak in one or more science content areas, while at the same time, existing alternative certification requirements are thought to be so rigid and time consuming that content experts, such as scientists and engineers, are discouraged from transitioning into teaching.

**Alternative Certification Programs and Evolving Online Experiences**

Zeichner and Schulte (2001) have defined an alternative teacher certification program as any alternative to a 4-year or 5-year undergraduate teacher education program including programs with lesser standards for certification as well as those that require certification candidates to meet the same standards stated by undergraduate teacher education programs at higher education institutions. These authors note that programs with lesser standards might more appropriately be called “recruitment” rather than “certification” programs. Considering “standards” that are appropriate for beginning teachers, Zientek (2007) has listed criteria for teacher preparation admission that include: knowledge of content and pedagogy, successful performance on certification examinations, participation in mentoring experiences, demonstration of verbal ability and belief in the value of education and teaching.

Harrell and Harris (2006) provide greater specificity in describing alternative teacher certification programs. In Texas, alternative programs are an array of post-baccalaureate programs that are offered through colleges, universities, school districts, the state education agency, educational service centers, community colleges and private providers. The curricular experiences for these programs are typically framed by time-concentrated teacher survival skill training prior to the candidate’s entry into the classroom as the teacher of record. During the first year of teaching, recurring mentoring by program and school personnel are provided to the candidate in order to become acclimated to classroom duties and responsibilities. Programs offered by colleges and universities typically admit candidates who meet admission requirements either for graduate school or admission into a teacher education program as well as demonstrate content mastery through passing the state content licensure examination. Successful completion of the program occurs when the candidate finishes the yearlong teaching experience receiving acceptable performance evaluations from program officials and passes the state’s Pedagogy and Professional Responsibilities (PPR) licensure examination.

Literature reviews on teacher preparation via distance learning yielded one online program offered at the University of North Texas (Harrell & Harris, 2006). Other listings from these reviews provide comparisons between an online course experience and an in-class course experience (Smith, Smith & Boone, 2000) and accounts of how web-assisted course experiences have augmented in-class teacher preparation courses.

**Accelerate Online/OPTIONS Teacher Certification Program**

This program was established to provide a flexible alternative certification program for life science and physical science, including chemistry, physics, composite science, and mathematics (grades 8-12). This program has three features setting it apart from other alternative certification programs offered in Texas. First, it is offered as a university-based continuing education program that does NOT yield student credit hours to the University, thus reducing costs (no tuition expenses) for candidates although they do pay an enrollment fee to participate in the
program. Second, because the pedagogy content associated with state licensure is accessible 24/7 as an on-line experience throughout the state, certification can be completed by those possessing or pursuing science degrees with an online program of education that can be completed in 12-18 months. Third, the program has been developed from a partnership between a College of Education and Human Development and a College of Medicine that is providing a talent pool of candidates with strong academic backgrounds in science and mathematics. The curricular elements of the program consist of an on-line curriculum with web-supported field experiences. The online curriculum contains 35 online modules developed to engage the candidate with concepts and skills identified as necessary for a beginning teacher by the Texas Board of Educator Certification. The Field Based Experience consists of a 40 clock-hour supervised teaching field experience in a secondary school during the initial phase of the program. The final program component is a paid internship, where candidates are supported by both a trained classroom teacher mentor and the university supervisor, who guide, observe and provide constructive feedback to the interns during their year-long development as a beginning teacher. These program components are described in greater detail elsewhere by Denton, Davis, Smith, Beason, and Strader (2005).

Recruitment & Selection of Participants

Candidate Recruitment. We have sought opportunities to personally connect with potential applicants through vis-à-vis interactions at career fairs, booths at professional conferences and personal visits with college and departmental academic advisors. These approaches, although quite time intensive, are yielding personal contacts leading to program candidates. Additional techniques we have implemented and continue to use to communicate our program to the educational community include: poster placements on bulletin boards at strategic locations on campus (close proximity to college advising offices); direct communication (e-mail, telephone, cover letter and brochure) with school district certification offices and secondary school principals; program announcements provided on an intranet bulletin board; program description links on college homepage and recently Google search engine “Ad words” and URL advertising based on keyword selection by potential applicants.

Gradually, these marketing efforts and personal contacts have succeeded in raising awareness about our program to the extent that 70 to 100 inquiries are received each month. During the past year, we registered 973 inquiries about the program. The cost of marketing the program to achieve this increased awareness has also increased proportionately with $2,000/month now being spent on recruitment.

Admission of Candidates. The requirements for admission into Accelerate Online/OPTIONS are based on admission requirements of the undergraduate teacher preparation programs at our University. In order to be admitted the applicant must:

• either be a college graduate or graduate student or be within one year of completing a baccalaureate. Applicants must have at least a 2.5 GPR in their baccalaureate course work attempted or a 2.75 GPR on the last 60 hours of the baccalaureate degree or 3.00 GPR on advanced degree course work.
• have a minimum of 24 semester hours in a specific content specialization or 36 semester hours in a composite field (including a minimum of 3 semester hours in all sub-areas).
• pass or be exempt from TASP (basic competency test). Holding a baccalaureate degree exempts the applicant from this requirement.
• pass the required Texas Examinations of Educator Standards (TExES) content exam(s). This requirement was included to enable our candidates to be considered “fully qualified” in a content area as defined by the “No Child Left Behind” guidelines.
• have access to an Internet-capable computer.
• pass a background check of any criminal activity relating to children.
• submit two letters of recommendation.
• demonstrate verbal fluency with English and convey a genuine interest for teaching youth in a personal interview with program staff.

Texas Examinations of Educator Standards (TExES) pedagogy and professional responsibilities and content examinations are copyright protected tests of the Texas Education Agency that have been prepared and distributed under contract with the Educational Testing Service. Content examinations are provided for each certification area with test administrations governed by regulations and standards of the Texas Education Agency. The purpose of these criterion-referenced tests is to measure the requisite knowledge and skills that an entry level teacher in the
certification area is required to possess (State Board for Educator Certification/Texas Education Agency [SBEC/TEA], 2006a-d).

Because the preparation of secondary mathematics and science teachers is the primary goal of *Accelerate Online/OPTIONS*, the following descriptions briefly present the test parameters for Pedagogy and Professional Responsibilities 8 - 12, Mathematics 8 – 12, Life Science 8 – 12, Physical Science 8 - 12. Chemistry and physics are components of the physical science examination. These examinations are composed of 90 multiple choice items with 80 scorable items and 10 items used for pilot testing that do not contribute to the examinee’s score. The examination is administered as a day long experience with a half-session in the morning and another in the afternoon. Each session is five hours long. The content domain and test item structure for these tests are:

- **Pedagogy and Professional Responsibilities 8 - 12** – designing instruction and assessment (25 items), creating a productive classroom environment (12 items), implementing effective instruction and assessment (25 items), fulfilling professional roles and responsibilities (18 items) (SBEC/TEA, 2006a);
- **Mathematics 8 - 12** – number concepts (11 items), patterns and algebra (27 items), geometry and measurement (15 items), probability and statistics (11 items), mathematical processes and perspectives (8 items), mathematical learning, instruction and assessment (8 items) (SBEC/TEA, 2006b);
- **Physical Science 8 - 12** – scientific inquiry and processes (11 items), physics (29 items), chemistry (33 items), science learning, instruction and assessment (7 items) (SBEC/TEA, 2006c);
- **Life Science 8 - 12** – scientific inquiry and processes (12 items), cell structure and processes (16 items), heredity and evolution of life (16 items), diversity of life (16 items), interdependence of life and environmental systems (12 items), science learning, instruction and assessment (8 items) (SBEC/TEA, 2006d);

These tests are scored on a 100 – 300 scale, with a criterion score of 240 needed to pass each of these examinations. State-wide results provided by the Texas Education Agency of initial pass ratios from the 2005-06 academic year for these certification areas are: pedagogy and professional responsibilities 87% (4882 passers/5599 takers), mathematics 62% (745 passers/1209 takers); life science 53% (310 passers/583 takers); physical science 71% (95 passers/133 takers) (Texas P-16 Public Education Information Resource, n.d.).

Demonstrating verbal fluency in English in a personal interview was added to the selection criteria after a number of candidates exhibited some difficulty in communicating orally. While the primary function of the interview is to determine oral fluency, this interview also provides the candidate an opportunity to meet project staff and seek additional information about the program before making a commitment. The project staff also assesses the applicant’s “soft” characteristics noted by Allen (2003) of high achievement orientation, accepting responsibility, demonstrating critical thinking, being organized, being motivated, being respectful of others and supporting the goals of the organization. Although rare, two applicants have been advised to seek other employment opportunities due to their interview responses.

*Candidate Program Costs.* For a candidate to begin the program, an Enrollment Agreement is entered into by the candidate and *Accelerate Online/OPTIONS*. This agreement delineates the cost of the program ($4,200) with different payment options, the availability of scholarships and the expected time to complete the program with provisions for extending the time to complete the program. Generally, candidates opt for an extended payment schedule with the convenience of online payments.

*Completing Online Curriculum and Early Field Experience*

*Monitoring Candidate Progress.* An extensive digital monitoring system was developed for *Accelerate Online/OPTIONS* that includes an on-line registration system with password protection for candidate entry and an underlying management tracking resource. For the candidates, the management system serves to affirm their program status in terms of completed assignments and module deliverables, and it will return them to the point they exited when they come back to the module. As a management resource for program administrators, this digital monitoring system attends to each candidate’s visits to a module, the elapsed time spent examining the contents of the module and whether items requiring a response have in fact been completed. Given these data, candidate progress in completing the modules are continually monitored to determine whether individual candidates are progressing satisfactorily in the program, as well as determining whether modules have potential design flaws given the collective performance of the candidates on particular activities and their overall performance on the module.
Early field experience and postings to ePortfolio. Early field experience activities are assigned to engage candidates in classroom actions in terms of the principles and concepts presented in the online modules. This field component has been designed for the candidate to experience a gradual induction into the teaching environment through observing quality teaching and gaining insights about the school’s organizational culture. The activities for this experience include: interviews, structured observations of effective teaching practices, a follow-me-teach activity where the candidate observes a teacher deliver a lesson to one period and then teaches another class the same lesson. This experience culminates with the candidate developing and teaching her own lesson to both classes. Across the multiple week experience, candidates are instructed to upload observations, interview forms, completed lesson plans and teaching reflections to their ePortfolio as documentation.

Candidate Placement and Support in Final Field Experience

Establishing effective strategies for placing candidates in paid teaching assignments has evolved from denial (affirming among ourselves that job placement was NOT our responsibility) to affirmation (actively marketing all of our candidates to school districts). Ultimately, our goal is to place all of our candidates that have completed their online modules and early field experiences in paid internships. As placement protocols have evolved, we have actively communicated the qualities of the Accelerate Online/OPTIONS program and the credentials of our candidates to school officials in assisting our candidates to obtain paid internships.

We have found that powerful tools in recruiting schools to hire our candidates are the ePortfolios the candidates have developed after beginning their programs. To facilitate the active review of the credentials of our candidates, an online ePortfolio Center http://empowermentzone.tamu.edu/portfolios/center/ was set up for school administrators to search ePortfolios of teaching candidates – by Last Name, Certification Area or the candidate’s preferred location or Regional Preference. Over the course of these events and activities, we have learned that contacting secondary school principals and human resource officials about our available candidates and providing them access to the candidates’ ePortfolios have been very effective placement tools.

Guidelines and responsibilities for the intern, the university supervisor and the mentor teacher are provided to each intern and their support team members just prior to the final field experience. These materials and activities reflect successful practices gleaned over time by teacher educators and university supervisors. To illustrate, interns are instructed to submit weekly instructional plans to their ePortfolio for review by their supervisors and mentor teachers in preparation for the coming week. Further, we have applied a strategy recommended by Morgan and Kristonis (2008) in hiring highly regarded retired teachers and principals as university supervisors. These experienced educators provide valuable guidance and support to the intern with their introduction to the formal and informal rules of the school and classroom. To illustrate their dedication to developing interns in becoming capable teachers, supervisors have assisted in developing a digital learning community for sharing ideas and digital resources. This learning community arose to address the need to discuss recurring challenges faced in supporting their interns.

Methods

Sample

The sample includes 170 individuals that have provided applications to the Accelerate Online/OPTIONS Secondary Teacher Certification Program. Table 1 presents descriptions of academic variables examined in this study. The academic variables selected are based on the program’s admission criteria as well as recommendations from the literature on recruiting and preparing quality teachers (Alberts, 1999; Allen, 2003; Darling-Hammond, 2006; Goldhaber & Brewer, 2000; NCTQ, n.d.; USDE, 2000; Zientek, 2007). The study consisted of 126 females and 44 males across five teaching fields: Life Sciences ($n = 95$), Physical Sciences ($n = 13$), Composite Science ($n = 22$), Mathematics ($n = 18$), Other ($n = 22$); holding PhD/MD ($n = 28$), MS ($n = 19$), or BS ($n = 123$) degrees. Five ethnicities were represented: Euro-American ($n = 121$), African-American ($n = 6$), Asian-American ($n = 9$), Hispanic ($n = 13$), Indian ($n = 9$), not specified ($n = 12$), and career status, either a recent graduate ($n = 108$), or a mid career changer ($n = 62$).
Table 1. Frequencies and Descriptive Statistics of Select Academic Variables

<table>
<thead>
<tr>
<th>Academic Variable</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduate Grade Point Ratio</td>
<td>144</td>
<td>3.12</td>
<td>.46</td>
<td>4.00</td>
<td>2.00</td>
</tr>
<tr>
<td>TExES Content Score</td>
<td>155</td>
<td>254.5</td>
<td>24.0</td>
<td>294</td>
<td>182</td>
</tr>
<tr>
<td>Semester hours in content field</td>
<td>83</td>
<td>43.5</td>
<td>13.6</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>Years since undergraduate degree</td>
<td>169</td>
<td>6.2</td>
<td>9.0</td>
<td>35</td>
<td>0</td>
</tr>
</tbody>
</table>

Summarizing the values in the preceding description and table as an applicant profile, we note that the typical applicant to our program is a Euro-American female who majored in a life science program accumulating more than 43 credit hours in their major and recently earned a baccalaureate degree while maintaining a B grade average. This profile is consistent with gender and ethnicity values reported by Zientek (2007) and supports the view noted by Good, et al. (2006) that alternative certification programs attract more teachers with degrees in science than undergraduate teacher preparation programs.

Data Collection

Extensive biographic data on applicants are compiled into a program data base beginning with the submission of an application. Assuming all admission criteria have been met, the applicant is then officially accepted into the program as a candidate. The final steps to be certified as a secondary science or mathematics teacher, assuming the completion of a successful final field experience, are that the candidate passes the TExES Pedagogy and Professional Responsibilities (grades 8-12) examination (i.e., meeting or exceeding the State’s criterion score of 240) and submits an application to program staff who verify the candidate’s successful completion of certification requirements. At present, every Accelerate Online/OPTIONS intern who has taken the TExES Pedagogy and Professional Responsibilities examination has attained a passing score. For comparison state-wide data obtained from the State Board of Educator Certification for initial pass ratios for the 2005 - 06 academic year for Pedagogy and Professional Responsibilities 8 – 12 test was 87% (4882 passers/5599 takers) (Texas P-16 Public Education Information Resource, n.d.). Scores from the TExES tests and additional program data are added to the record of each candidate as the information is received resulting in a record containing more than 300 variables on each candidate that successfully completes the program.

Data Analysis

The Statistical Package for the Social Sciences was used to conduct descriptive statistical analyses of the biographic and academic variables across this sample to determine the status of applicants with respect to completing teacher certification in our program. The resulting summaries presented in the preceding Sample section confirm sufficient diversity across the applicants to warrant additional analyses of these data. Subsequently, non-parametric correlations were calculated to determine whether particular profile variables of applicants were related to successful program entry and program completion. Discriminant analyses were then conducted to determine whether biographic and academic profiles can be identified to predict both successful program admission and completion to become a certified secondary teacher.

Results

Descriptive Analyses

Participant status has been monitored as an important formative evaluation function in determining our progress in implementing the program. Table 2 provides a summary of participant status across three years of program operation. We note from examining the values in Table 2, a sizeable number of applicants across all levels of academic preparation were not admitted into our online certification program. Attrition ratios (# of applicants not admitted/ # of applicants) partitioned by highest degree attained were found to be .57 for PhD/MD, .42 for MS and .39 for BS applicants. Reviewing admission criteria and corresponding data, four factors were found that inform non-admission decisions. These factors and the number of applicants affected are presented in Table 3.
Table 2. Status of Individuals Submitting Applications to Accelerate Online/OPTIONS
Program Partitioned by Highest Degree Earned

<table>
<thead>
<tr>
<th>Highest Academic Degree</th>
<th>PhD/MD</th>
<th>MS</th>
<th>BS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Admitted</td>
<td>16</td>
<td>8</td>
<td>48</td>
<td>72</td>
</tr>
<tr>
<td>Current Applicant</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Dropped Candidate</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Current Candidate</td>
<td>4</td>
<td>4</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Current Intern</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Dropped Internship</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Completed Certification</td>
<td>7</td>
<td>5</td>
<td>50</td>
<td>62</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>28</td>
<td>19</td>
<td>123</td>
<td>170</td>
</tr>
</tbody>
</table>

Table 3. Reasons for Program Non-admission of Accelerate Online Applicants

<table>
<thead>
<tr>
<th>Highest Degree Attained</th>
<th>PhD/MD</th>
<th>MS</th>
<th>BS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not admitted</td>
<td>16</td>
<td>8</td>
<td>48</td>
<td>72</td>
</tr>
<tr>
<td>Failed TEES Content Test</td>
<td>6</td>
<td>0</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>Unsuccessful Interview</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Low GPA</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Incomplete Application</td>
<td>9</td>
<td>7</td>
<td>21</td>
<td>37</td>
</tr>
</tbody>
</table>

The common reason for non-admission has been failing to complete the application process. Thirty-seven (37) applicants have failed to provide one or more required elements needed for a complete application to the program and have not been admitted. Apparently a “change of heart” about becoming a teacher occurred for the applicant during the process of submitting a complete application packet. Possibly the tasks of obtaining transcripts to document their GPA and course work in their specialization, the background check of criminal activity, obtaining two support letters and the prospect of not successfully completing the TEES content test caused some applicants to reconsider their career decision about becoming a teacher. Perhaps these admission requirements were so rigid and time consuming that scientists, engineers and some recent graduates with science majors were discouraged from transitioning into teaching as suggested by recent reports (NRC 2000b; USDE, 2002). Yet these admission procedures are thought to be necessary to assure ourselves and employing school districts the applicant has been carefully screened to teach secondary students and exhibits characteristics considered important by teacher educators (Zientek, 2007).

Low undergraduate grade point ratios resulted in non-admit decisions for 5 applicants with BS degrees, although each of these individuals had transcripts that documented sufficient academic courses in the teaching field to meet admission requirements. Less pronounced but meaningful is the listing of two applicants who were not being admitted because personal views and expectations about teaching were expressed during their interviews that are quite incongruent with actual classroom environments. The soft attributes cited previously (Allen, 2003; NCTQ, n.d.) about personal attributes of high achievement orientation, accepting responsibility, demonstrating critical thinking, being organized, being motivated, being respectful of others and supporting the goals of the organization have been integrated into the interview process and assessed on each applicant.

Twenty-eight (28) of the 72 applicants (39%) that were not admitted to the program did not attain the TEES content test criterion score for admission. This percent of unsuccessful test takers correspond to unsuccessful test takers in mathematics (38%) and life science (47%) state-wide (Texas P-16 Public Education Information Resource, n.d.). Masters level applicants performed as expected with all of these applicants reaching or exceeding the criterion score of 240. Yet similar results did not occur for the PhD/MD applicants with 6 applicants holding advanced degrees not achieving the criterion score. Also, 22 applicants holding baccalaureate degrees with academic concentrations in the tested content domain failed to reach the admission criterion score on the TEES content test. Although reasons for the less than expected performance of PhD/MD applicants have not been determined, one conjecture is that the advanced degree applicants have become so specialized in their graduate or professional studies that some areas of the content domain addressed on the test had not been considered for several years. A second hypothesis is that these individuals over-analyzed the test questions. Whatever the reason or reasons for these less than expected performances, this TEES criterion score standard for admission has impacted whether an applicant has become a
candidate in our teacher preparation program. This finding suggests the content test admission standards are not as lax as has been suggested in the literature (USDE, 2002).

Looking at the program participants from the perspective of program completers was then undertaken. In the Table 4 we examined the ratios of success for attaining certification.

<table>
<thead>
<tr>
<th>Highest Degree Attained</th>
<th>PhD/MD</th>
<th>MS</th>
<th>BS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Completing Certification (A)</td>
<td>7</td>
<td>5</td>
<td>50</td>
<td>62</td>
</tr>
<tr>
<td>Number of Applicants (B)</td>
<td>28</td>
<td>19</td>
<td>123</td>
<td>170</td>
</tr>
<tr>
<td>Success Ratio (A/B)</td>
<td>0.25</td>
<td>0.26</td>
<td>0.41</td>
<td></td>
</tr>
</tbody>
</table>

With these ratios clustering at .31 +/- .10 when partitioned by highest degree of applicant, other biographic variables (gender, ethnicity and career status) were examined to determine whether an observed relation occurred when individuals that successfully completed the program and were certified were partitioned by these biographic variables. The results of these analyses are presented in Table 5.

<table>
<thead>
<tr>
<th>Applicants</th>
<th>Gender</th>
<th>Ethnicity</th>
<th>Career Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attained Certification (A)</td>
<td>Female</td>
<td>Euro-American</td>
<td>Recent Graduate</td>
</tr>
<tr>
<td>Total Applicants (B)</td>
<td>126</td>
<td>109</td>
<td>108</td>
</tr>
<tr>
<td>Ratio (A/B)</td>
<td>.39</td>
<td>.46</td>
<td>.44</td>
</tr>
</tbody>
</table>

The ratios for gender, ethnicity and career status cluster at .35 +/- .11; it appears these biographic variables are modestly associated with the applicant ultimately completing the program and becoming a certified secondary teacher. Yet, if we assume the five current interns (see Table 2) who are female, Euro-American and recent graduates will complete certification, then variation across these ratios increase and lead us to examine these data from another perspective.

**Bivariate Correlations**

Non-parametric correlations were calculated between the variable, program admission with two nominal academic variables (criterion TExES score, criterion undergraduate GPA), one ordinal variable (highest degree) and three nominal biographic variables (Euro-American, gender, career status), and then repeated for the variable, completed certification with those same academic/biographic variables.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Admitted</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>r</td>
<td>Sig.</td>
</tr>
<tr>
<td>Criterion TExES Content Score</td>
<td>155</td>
<td>.612</td>
</tr>
<tr>
<td>Criterion Undergrad. GPA</td>
<td>144</td>
<td>.106</td>
</tr>
<tr>
<td>Highest Degree</td>
<td>170</td>
<td>-.102</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>144</td>
<td>-.163</td>
</tr>
<tr>
<td>Gender</td>
<td>170</td>
<td>-.069</td>
</tr>
<tr>
<td>Career Status</td>
<td>167</td>
<td>-.213</td>
</tr>
</tbody>
</table>

Statistically significant correlations were found between becoming a candidate with attaining criterion TExES content scores ($r = .612$) and career status ($r = -.213$) favoring recent graduates becoming candidates. Similarly, statistically significant correlations were found between becoming certified with attaining criterion TExES content scores ($r = .392$) and career status ($r = -.231$) favoring recent graduates. The relation between program admission and
meeting or exceeding the criterion score on the TExES content test was expected because a TExES criterion score must be reached before the applicant can be admitted and ultimately certified; in fact, Table 3 provides values to determine that 39% of the applicants not admitted did not reach the TExES criterion content score. The negative correlations for career status occurred due to coding the recent graduate as “0” and the mid-career applicant as “1”. Given the values reported in Table 5 for candidates attaining certification, it was expected that more candidates and program completers would occur under “recent graduates” than “mid-career,” and the negative correlations support these expectations. The biographic variable ethnicity presented as a nominal variable, Euro-American (coded as 1) and Other (coded as 0), yielded a correlation ($r = .163$) with being admitted as a candidate and a correlation ($r = .151$) with becoming a certified secondary teacher. These values align with the success ratios presented in Table 5 where the Euro-American ratio exceeded the Other ratio in terms of the ethnicity of candidates completing certification. Finally, the academic variable, highest degree presented as an ordinal variable with doctorate (coded as 3), masters (coded as 2) and baccalaureate (coded as 1) yielded a correlation ($r = -.136$) with completing teacher certification. This correlation, although not statistically significant, is consistent with the values presented in Table 4 where a higher success ratio and a substantially greater number of baccalaureate candidates completed certification compared to candidates holding masters and doctoral degrees.

**Discriminant Analyses**

A final set of analyses was then conducted to determine whether dichotomous biographic variables (i.e., ethnicity [Euro-American or Other], career status [mid-career or recent graduate]) and academic characteristics of applicants (i.e., criterion TExES content score attained [yes/no], highest degree [3 = doctorate, 2 = masters, 1 = baccalaureate]) could be combined to predict whether the applicant would become a candidate in our program and complete certification. These biographic and academic variables, presented in Table 6, were selected for these analyses based on the magnitude of the correlations between these variables with becoming a candidate and completing certification. Discriminant analysis was selected because this procedure is useful for situations where a predictive model of group membership is desired based on observed characteristics of each applicant. The procedure produces a single discriminant function based on linear combinations of the predictor variables that provide the best discrimination between the two groups, (i.e., candidate/non-candidate or certified/not certified). The function is generated from a sample of cases for which group membership is known; the function can then be applied to new cases with measurements for the predictor variables but unknown group membership (Stevens, 2002).

Table 7 summarizes two discriminant analyses conducted for classifying whether applicants would become candidates in the program and then complete certification. Each of the two canonical discriminant functions yielded the following values: canonical correlation, Wilks' lambda, standardized function coefficients and classification results using cross validation of the left one out method (Huberty, 1994).

<table>
<thead>
<tr>
<th>Table 7. Summary of Canonical Discriminant Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discriminate Function</strong></td>
</tr>
<tr>
<td><strong>Standardized Coefficients</strong></td>
</tr>
<tr>
<td>Program Candidate</td>
</tr>
<tr>
<td>TExES content criterion score</td>
</tr>
<tr>
<td>Ethnicity</td>
</tr>
<tr>
<td>Career Status</td>
</tr>
<tr>
<td>Highest degree</td>
</tr>
<tr>
<td>Group Centroids (yes)</td>
</tr>
<tr>
<td>(no)</td>
</tr>
<tr>
<td>Correctly Classified Cases</td>
</tr>
<tr>
<td>Correct Predicted</td>
</tr>
<tr>
<td>Incorrect Predicted</td>
</tr>
<tr>
<td>Sample size</td>
</tr>
<tr>
<td>Canonical Correlation</td>
</tr>
<tr>
<td>Wilks' Lambda</td>
</tr>
<tr>
<td>Statistical sig.</td>
</tr>
</tbody>
</table>

*Note. *Variable omitted from analysis
For each discriminant analysis, the sample size was reduced from values reported previously (see Table 1) due to missing data, usually the TExES content score. Results of the discriminant analyses revealed that each function was statistically significant ($p < .001$) with Wilk’s lambda values being .516 for admission into program and .584 for completing certification. Because each analysis consisted of just two groups, i.e. candidate/non-candidate or certified/not certified, only one discriminant function per analysis could occur (Stevens, 2002). These analyses also reveal that these discriminant functions do distinguish group membership reasonably well. The canonical correlation between the discriminant function and program admission was .696. For becoming certified, it was .645. Further, the distance between the group centroids for the two functions ranged from 1.67 to 2.07 with the function exhibiting the greater inter-centroid distance yielding a greater number of correct group classifications.

The standardized discriminant function coefficients reported in Table 7 indicate the relative importance of each of the independent variables in classifying program admission as a candidate and program completion or certification when all of the predictors were added at once. For the discriminant function for program admission, the TExES content criterion score variable provided the largest coefficient in determining group membership, while for the discriminant function predicting certification, the largest coefficient occurred for the variable program candidate. The relative importance of the other predictors varied depending on the focus of the discriminant function.

The percentage of cases correctly classified or “hit rate” was 86.7% for program admission and 78.1% for completing teacher certification. The case wise statistics (discriminant scores, actual group and predicted group membership per individual) for each analysis combined with an examination of the raw data file revealed that for the program admission analysis, all of the misclassified cases occurred due to the applicant electing to withdraw from the program for reasons that are not captured by the four independent variables applied in this analysis or any variable in our program data base. Similarly, all of the misclassified cases for the completing teacher certification analysis occurred because these individuals are currently “in-progress” as candidates or interns.

**Conclusions**

Applying the results of the canonical discriminant function - admitted to program to Research Question 1, the TExES content criterion score was found to be the most influential variable. The relative importance of the other predictors were quite modest with highest degree providing some influence and career status and ethnicity providing very limited influence on predicting whether an individual would become a candidate. The values reported in Tables 3 – 6 for the analyses leading to the discriminant analysis confirm the relative importance of the TExES content criterion score in predicting whether the applicant would become a candidate in the Accelerate Online/OPTIONS program. To illustrate, in Table 3, we note that 39% (28/72) of the applicants failed to attain the criterion score and were not being admitted to the program. Table 6 contains the non-parametric correlation between TExES content criterion score and program admission exhibiting the greatest magnitude among the correlations reported. Of the remaining predictors in the discriminant function, only career status yielded a statistically significant correlation with program admission. The other predictors, highest degree and ethnicity did not yield statistically significant correlations with becoming a candidate, but did account for variation in the success ratios of applicants becoming certified (Tables 4 & 5). Other academic variables (undergraduate grade point ratio, accumulated semester hours in teaching field) often cited in the literature (Allen, 2003; NCTQ, n.d.; USDE, 2000) and biographic variables (gender and years following BS degree) were considered, but not included in the final set of analyses due to small quantitative associations with program admission. The low correlations for cumulative undergraduate grade point ratios with program admission occurred because our criterion for this admission variable includes minimum values that restricted the range of values for this variable across the applicants.

As program developers, we are pleased that the biographic variables (gender, ethnicity and career status) yielded little quantitative support to indicate these variables significantly influence admission decisions into our certification program. If these variables had been found to influence admission decisions, we would have had immediate adjustments to make in our recruitment and admission protocols to eliminate discriminating against applicants on the basis of gender, ethnicity and age (career status). However, the observation from our analyses that 42% (72 of 170) of our applicants do not become candidates indicates our recruitment and admission procedures need refinement. The most common reason for not becoming a candidate is failing to complete the application process suggesting the applicant has a “change of heart” about becoming a teacher. Perhaps more attention to the “soft attributes” of a teacher is needed in our recruitment materials for the applicant to consider before initiating the application process.
Other factors may be program costs (time and financial) and lack of flexibility in meeting admission requirements. In terms of costs, both in terms of dollars ($4,200) and time to completion (6 months to enter classroom as the teacher of record, and an additional 12 months to complete a yearlong paid internship and certification) are comparable to other alternative certification programs in the region and less expensive in time and money than a university program leading to certification and earned credit toward a graduate degree. As noted previously, obtaining transcripts to document grade point averages and course work, undergoing a background check for criminal activity, obtaining support letters and successfully completing the TExES content test may cause some applicants to reconsider their career decision about becoming a teacher. Perhaps these admission requirements are considered to be so rigid that prospective teaching candidates are discouraged from transitioning into teaching as suggested by recent reports (NRC 2000b; USDE, 2002). Yet we believe these admission procedures are necessary to affirm the candidate’s fitness to teach in secondary schools.

In formulating a response for Research Question 2, we again examined the results of a discriminant analysis, in this case the canonical discriminant function for completes certification, and found the most influential variable to be whether the individual had become an active candidate in the program. Of the other predictors, career status provided some influence in predicting whether a individual would complete certification, but the other variables’ predictive power were subsumed into the variable, active candidate. Because this discriminant function on predicting program completion is sensitive to the status of active candidates, given that all of the misclassifications of this function were active candidates, perhaps this function should be re-analyzed next year after program completion data of active candidates and interns will be more complete.

However, the status of individuals submitting applications reported in Table 2 support the generalization that once an individual becomes a candidate, she will very likely continue in the program and complete certification. At present, just 6 of 98 candidates or former candidates (6%) have elected to withdraw from the program. Exit interviews with interns that withdrew from the program suggest greater attention to the soft attributes during the recruitment and application processes may prevent a change of heart about pursuing a teaching career. It seems that for these disillusioned interns, idealized expectations held about secondary teaching were immediately challenged. Expectations about guiding and nurturing highly motivated students who were inherently interested in the course content were called to question the first day in the classroom. Facing more than 100 students each day, many of whom appeared unmotivated and disinterested, immediately produced a sense of being overwhelmed and frustrated that shortly turned to despair. Acting upon specific recommendations of supervisors and applying the tenets of curriculum organization they had experienced in the online curriculum were not cognitively possible given the disequilibrium they were experiencing each day. It is quite likely unrealistic ideas about teaching and students led these beginning interns to the decision to exit the classroom. Perhaps the untenable circumstances these ill-fated interns found themselves in could have been prevented had a more formal process been in place to assess their unsupported beliefs related to: assuming high achievement orientation, accepting responsibility, demonstrating critical thinking, being organized, being motivated, being respectful of others and supporting the goals of the organization.

In summary, we have found from this investigation that attaining a criterion score on the TExES content test is the best single predictor for determining whether an applicant will become a candidate in our program. This finding reinforces the intensive diagnostic effort our admissions officer invests in reviewing applicant transcripts, administering and scoring a practice certification test and providing feedback to the applicant on their probable success with/without additional remediation before sitting for the state’s content certification test. This support is provided to applicants as they consider whether to invest the time and money to become a secondary teacher. Further, if the applicant matriculates into the program, then the best variable to predict completing certification is the variable, program candidate. While the various statistical analyses conducted have produced internally consistent results, we have found that we may be able to further inform our recruitment, selection and placement protocols by incorporating inventories that seek to measure the applicant’s soft attributes thought to impact whether the beginning teacher succeeds and completes their internship and certification.

Acknowledgements

Support provided by U. S. Department of Education, Transition To Teach Program and the Houston Endowment Foundation, Incorporated.
References


A Distributed Online Curriculum and Courseware Development Model

Pinar Onay Durdu
Kocaeli University, Department of Computer Engineering, 41040, Kocaeli, Turkey // pinar.onaydurdu@kocaeli.edu.tr

Nese Yalabik
Middle East Technical University, Department of Computer Engineering, 06531, Ankara, Turkey // yalabik@ceng.metu.edu.tr

Kursat Cagiltay
Middle East Technical University, Department of Computer Education and Instructional Technology, 06531, Ankara, Turkey // kursat@metu.edu.tr

ABSTRACT
A distributed online curriculum and courseware development model (DONC²) is developed and tested in this study. Courseware development teams which may work in different institutions who need to develop high quality, reduced cost, on time products will be the users of DONC². The related features from the disciplines of instructional design and software engineering were combined. The research is conducted as a collective case study, including four cases with distinctive characteristics to reveal the several practices in online curriculum and courseware development work. The DONC² development model was proposed using the results gathered from the investigated cases and a literature survey. The model uses the iterative incremental and agile software development approaches in order to overcome the disadvantages of other linear development approaches. This enables building releasable yet good quality products in short time periods. Furthermore, continuous communication, evaluation and feedback as well as good project management and readiness to adapt to changes are integrated as the essential characteristics. DONC² is different then previous linear and non-adaptive models in all of these aspects. The model was tested with one of the cases while being improved with success. It was applied as a development model for scheduling of the courseware development project in the last case and considered as helpful for the success of the project by the project team members.

Keywords
Virtual learning environments, Online courseware and curriculum development, E-learning, Adaptive software development

Introduction

The number of people who are seeking a university degree, skill enhancements or lifelong learning has increased tremendously. This has forced universities and companies to find new ways to provide education to the mass learners and recent developments in information technology and Internet have enabled that by delivering web-based courses via “virtual learning environments” (VLE) (Xu, Wang & Wang, 2005, p.525). Many institutions have started projects to employ e-learning, which has the goal of “learning anytime and from any place” (Barjis, 2003, p.4).

VLEs are the online learning systems which provide a complete learning environment including various features such as course materials, evaluation instruments or communication and collaboration tools (Ryan, Scott, Freeman & Patel, 2000). They can be considered as the new form of providing education. Their features should be very different from the traditional classroom settings since there is no face to face interaction between the instructor and the students. On the other hand, VLEs may provide many additional opportunities for achieving enhanced and enriched learning outcomes through the use of the web for effective instruction and can be a promising alternative to traditional settings (Zhang, Zhao, Zhou & Nunamaker, 2004). Web-based teaching can facilitate learner interactivity and also can provide a great amount of resources. In order to enhance learner interaction, courses can use innovative and dynamic learning materials. Unfortunately, these advantages are still not being used effectively since usually traditional instructional design methods are transferred to the web as if they were in the traditional settings. In addition, there is still no commonly accepted framework to guide developers in their design of curricula (Oliver & Mcloughlin, 1999). On the other hand, collaboration of a number of institutions might be important for the use of best resources in online material development and teaching.
For the success of VLEs, the effectiveness of courseware provided is essential. In order to provide effective courseware, careful consideration is needed during the development stages with the help of several disciplines such as instructional design, software engineering and human computer interaction (HCI).

ADDIE (analyze, design, develop, implement and evaluate) is a prescriptive sequential instructional design model, which describes the essential components of any instructional development process. Many other models that use this classical sequential approach have been developed. The “Instructional Design Plan” is a relatively new model for individual course development. It was developed by Kemp, Morrison and Ross (2004) and is different from ADDIE since it does not apply any specific sequence and considers the design and development process as a continuous cycle as can be seen in Figure 1. This model also covers issues like project management or implementation in addition to the instructional design steps.

![Diagram of Instructional Design Plan](image1)

**Figure 1.** Instructional Design Plan (Kemp, Morrison & Ross, 2004, p.1)

Some of the instructional design models are adapted to online course development (Tripp & Bichelmeyer, 1990, Willis, 1995, 2000). ‘Rapid prototyping’, proposed by Tripp and Bichelmeyer (1990) is an example. It does not use a sequential approach either. It is the best known model that makes use of the software engineering methodologies and the steps are realized almost in parallel with iterative development as seen in Figure 2.

![Diagram of Rapid Prototyping](image2)

**Figure 2.** Rapid prototyping approach adapted to Instructional Design (Tripp & Bichelmeyer, 1990)

Developing educational software has many common aspects with software development. Especially the design and production stages are similar since the product as well as the production medium and tools are the same for both.
Recently, software practices which are commercially proven to be successful in the software engineering field introduced iterative and agile methods. These focus on iterative development, managing requirements, using component based architectures, visually modeling software, verifying software quality and configuration management (Kruchten, 1998).

Agile methods call for lighter weight and faster development methods. Although these methods are sometimes considered as ad hoc or unstructured, they involve balanced structure and flexibility in their core, which is claimed to increase creativity and innovation (Highsmith, 2002). The important features that differentiate agile development from former models can be summarized as (Abrahamsson, Salo, Ronkainen & Warsta, 2002):

- The emphasis is on building releasable software in short time periods.
- The time period used for iterations is measured in weeks rather than months.
- They focus only on the functions required at first hand, delivering them fast, collecting feedback and reacting to received information.
- Close team relationships, close working environment arrangements and procedures that improve team spirit are at the core of the models.
- Continuously producing tested working software, collaboration with the customer and responding change.

Adaptive Software Development (ASD) (Highsmith, 2000) is an example framework among the agile approaches which emphasizes project management and collaboration practices (Highsmith, 2002). It is going to be explained in Section 2 in more detail.

Incorporating HCI, more specifically the usability issues to the development approach is essential for the overall quality of learning since the user will judge the system on the basis of interface which is the first contact point (Faulkner, 1998). Usability ensures the ease of learnability of the learning environment as well as the learning content, which increases the effectiveness. Therefore, in VLEs, since the core aim is to learn the contents rather than learn to use the system, the more usable the system is the more effective the learning takes place. Standard ISO 9241 defines usability as “allowing the user to execute his task effectively, efficiently and with satisfaction in the specified context of use” (Abran, Khelifi, Suryn, Seffah, 2003, p.331). In addition to that, usability approaches are also considered since they focus on formative and summative evaluation methods needed to be implemented throughout the whole development process in order to ensure the effectiveness of an instructional product (Crowther, Keller, Waddoups, 2004).

![Figure 3. Three Iterative UCD Phases (Detweiler, 2007, p. 41)](image-url)
Usability issues or concerns are not systematically covered in many software development approaches as there are no known methods to integrate these concepts to the development life cycle. HCI issues are generally considered only at the screen-interface or at the final design stages (Zhang, Carey, Te’eni & Tremaine, 2005). On the other hand, many of the software developing organizations are beginning to pay more attention to the usability of their products. In addition they also realize the importance of implementing these techniques early in the development processes (Ferre, 2003; Anderson, Fleek, Garrity & Drake, 2001). In a more recent study, Detweiler (2007) proposed three iterative phases which are repeated in all development phases of an agile software development project. This approach enables users to test the system from the beginning to the end of the development effort and to design, prototype and develop user interfaces iteratively, as seen in Figure 3.

The importance of usability is also realized in instructional development field. In addition to traditional evaluation done generally at the end of development cycle, the need to incorporate usability evaluation from the beginning to end is has been acknowledged by many instructional designers. Therefore, systematic formative and summative evaluation methods are needed to be implemented throughout the whole development process in order to ensure the effectiveness of an instructional product (Crowther, Keller & Waddoups, 2004).

Another important issue in the development of educational software is the use of related industry standards. Technology enhanced learning environments are requested to be compliant with current learning technology standards such as IEEE LTSC, AICC, and IMS (Anido, et al., 2002).

The literature review reveals that an integrated development model for distributed curriculum and courseware development for online learning environments has not been studied. In the present study, a collaborative online curriculum development methodology based on the disciplines discussed above (instructional design, software engineering, HCI) is proposed.

The ASD approach mentioned above is adapted as the software development methodology. ASD is described in detail in the next section. The design method used in this study, the outcomes of the investigated cases and the developed methodology and conclusions are presented in the subsequent parts.

**Adaptive Software Development (ASD) and Its Adaptation**

ASD is an iterative development model that stemmed from rapid application development. However, it is primarily different from them in that its acceptance of emergent order. It provides a dynamic speculate-collaborate-learn life cycle which is different from static plan-design-build life cycle models. It uses iterative cycles like spiral or evolutionary development but it also has another dimension which are represented by secondary arrows leading away from the iterative cycle, as can be seen in Figure 4. These indicate the emergent ideas that can be revealed throughout the development process. In complex environments it is usually not possible to determine all specifications at the beginning. Therefore speculation is offered as a replacement. Collaboration is required to balance unpredictable and predictable specifications. Learning occurs as a result of this collaboration. Stakeholders make small mistakes based on their false assumptions and they learn from their mistakes and gather better experience and mastery. As a result, this dynamic cycle enables continuous learning and adaptation to the emergent situations (Highsmith, 2002).

![Figure 4. The Adaptive Development Life Cycle (Highsmith, 2002, p.41)
The ASD emphasizes adaptation rather than optimization. It assumes that change and flexibility is necessary, so change management is at the core of the development methodology. It is also a component-based rather than a task-based approach and components are conducted in adaptive cycles throughout the project. The features of adaptive cycles involve the characteristics of ‘mission-focused, feature-based, iterative, time-boxed, risk-driven and change-tolerant’ (Highsmith, 2002, p.83). These short time-boxed iterative development cycles are primarily different than the sequential life cycles in that all life cycle activities are performed throughout the whole life cycle but with different workloads not only at corresponding phases as can be seen in Figure 5. Time-box representation looks similar to Gantt charts but it is actually different. It represents the amount effort for a particular task relative to other tasks. Wider the black area, more time/effort needs to be spent for the particular task. Since the horizontal axis represents time, the overlapped effort among tasks can also be observed. It also shows the amount of concurrency among tasks.

Change is considered as a likely event for online curriculum and courseware development effort so the model emphasizes continuous evaluation and revision. Importance attributed to continuous collaboration and communication is the other essential feature of the model. It does not provide a rigid structure for development efforts, instead, it focuses on flexibility and guides the developers how to conduct all these components. Since the ASD provides a framework which corresponds to these issues and emphasizes more on project management and collaboration practices, it is used as primarily adopted software development approach for the DONC2 model.

**Design of the Study**

An iterative incremental approach was also in the design of the study. In the first iteration, two online courseware development projects were investigated. At the end of this iteration a draft version of the development model was defined. In the next two iterations, validation activities for the proposed model took place and the final form of the model was established after the necessary revisions.

Online courseware development projects were examined as cases to see the problematic points as well as best practices that can be applied to the proposed development model. Some criteria which can be seen in Table 1 were determined for the selection of the cases but not all of them could be satisfied by all the cases.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum development focus</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>Online course material development</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Geographically distributed</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>development environment</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>Teamwork</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>
Data collection methods

Interviews were conducted with experts who participated in the projects examined with different roles, to learn about their experiences. In addition to the interviews, the courses developed in the project by different people were also examined to see whether the applied project practices lead to successful results.

Interviews

Three different interview question sets based on the nature of the investigated cases were used throughout the study. The first interview question set consisted of open-ended questions for the semi-structured interview sessions. In those interview sessions, the researcher tried to gather as much information as possible about the reactions of the team members for the courseware development processes. The interview questions were grouped into two categories. The first category involved the general questions, which tried to gather general information about the project as well as the interviewees. The second category involved main questions related to the development process, starting from the strategic decision making to the lowest level activities. These main questions can be classified into three layers.

- **Management layer**: This layer deals mainly with the Project management level strategic decision making activities as well as issues which are directly related with the project manager or project management team.
- **Integration layer**: This layer mainly acts as a middle layer among the management layer and the micro layer. It deals mainly with the curriculum development activities.
- **Micro level**: This layer is mainly deals with the courseware development level activities. It includes all issues to be considered for the development of an individual courseware.

The second interview question set was an evaluation matrix (EM) which was developed based on the proposed development model. This EM was applied to a developed courseware project as an evaluation framework in a structured interview session. There were four matrices covering the three layers of development processes as given above in addition to interlayer processes of communication and evaluation and revision. Matrices include all the elements that should be integrated to the development effort and check whether they were realized in the investigated project as well as at what level they were carried out.

The third interview set was again formed of open-ended questions for the semi-structured interview session. Here, the researcher tried to gather information about the developer’s views on the processes applied to the project.

Implementation Process

The implementations procedure and the investigated cases will be explained by giving background information about the subjects as well as the scope of the projects. The cases were studied in three iterations:

**Iteration 1**

In the first iteration two different instructional system development projects were examined.

Case 1- Avicenna Virtual Campus (AVC) Phase 1 (AVC-1): The major purpose of this project was to create a new community of universities from 14 Mediterranean countries by forming Avicenna Knowledge Centers (AKCs). Each AKC consists of a team of project director, pedagogical expert, technical expert and technicians and these teams has to develop about 20 online course modules in English and in their own language (Avicenna, 2006). The courses developed by the countries are uploaded to a content management platform called “Plei@d” for sharing. This project does not aim to develop any degree program. The phase I of the project was completed at the time of investigation.

Case 2 - SBS- Malaysia Project (SBS): This case was an e-learning project conducted by Siemens Business Services (SBS) for Malaysia Government. The major purpose of this project was to develop computer-aided materials for 7th and 8th grade mathematics and science courses of Malaysian formal elementary school education. It was a collaboration project of geographically distributed teams, which were located in Turkey and Malaysia. In total 330
courses were converted in about sixteen months period successfully. All of these courses have begun to be used by the Malaysian teachers and students at the schools since 2005 (Rtb, 2007).

Iteration 2

In this second iteration, Equipping Primary Care Physicians to Improve Care of Children (EPPICC) case was investigated. The first interview questions were used as in the previous two cases. In addition to this, additional interview sessions were conducted with the EM matrix developed based on the developed methodology.

Case 3 – EPPICC: The EPPICC is one of the online collaborative grant projects developed by the University Of Alabama Division Of CME (Continuing Medical Education) with the collaboration of different departments. It is jointly sponsored by the University of Alabama, Alabama Medicaid, and the National Institutes of Health. The goal of the EPPICC project is to provide online modules to improve the initial screening that pediatricians do with regard to eye care of small children (EPPICC, 2006). The modules are developed by the CME in collaboration with the Principal Investigators (PIs) of the grant. PIs were also responsible for the project management so can be considered as project managers. The CME provides courseware development team and PIs provided the content for the courses. The development of the modules was completed at the time of investigation.

Iteration 3

In this final iteration, the proposed methodology was applied as a development methodology to the Avicenna Virtual Campus Phase 3 project. The project plan was formulated with the appropriate components that are necessary for this phase of the Avicenna project. The project management component from the global layer; training and style guideline components from the macro layer and all micro layer components in addition to the communication and continuous evaluation and revision components were adapted for the project and implementation of the project was conducted based on this. After completion of the project, the third interview set was used.

Case 4 - AVC Phase 3 (AVC-3): In this case, the courseware development effort of the Turkish partner in the last phase of the AVC project was investigated. In this phase six more courses were developed in the scope of the project mentioned in Case 1 by the Turkish AKC. The same requirements were still used as in the previous phases.

Data Analysis

The data analysis was continuous and iterative throughout this work. The researcher first looked for similarities in the data from the participant interviews and the documents of the development effort. Secondly, the researcher looked for data that captured major differences among the efforts. Finally, categories and themes were created to arrive at a list of critical principles.

Outcomes of the Case Studies

The results of the interviews conducted with the team members of different projects and examination of their documents are summarized in the following section. Based on the answers of the interviewees, the major themes necessary for each layer, global, macro and micro layers have been determined.

Interview Results

Project managers, pedagogical experts or software developers who are responsible for strategic decision making level of their projects were interviewed throughout the study. Some principles were extracted and developed from the findings of the cases. These principles are applied in the various stages of the development process, and each has some (more or less) impact on different phases of the development model.
Outcomes of the AVC-1 Case

Providing a good project management team which is supportive, guiding and understanding is essential for the project, as they will enable the coordination among the distributed members is the most prominent principle gathered from this first case. Having people who have required competencies in the team is very essential for good decision-making. On the other hand, technical issues can be distracting during this kind of development projects and can demotivate the teams. Therefore, including a technical support team that will respond the technical needs of development teams in a timely manner is required for the smooth running of the project. Moreover, without adequate participation of any team members, the development approach cannot be succeeded. Training is required for the developers before and during the development process in order those to have the common understanding of the development process as well as what kind of material is going to be produced at the end. In this case, it was revealed that the capabilities of LMS were very limited and produced difficulties. The capabilities of LMS should be determined based on the requirements of the courses to be presented so the selection of LMS is an issue to be concerned during the development effort. Another issues revealed was that there is a need to apply an instructional design and strategies for the development of the courses as the materials are needed to have pedagogically sound principles to aid learning. Furthermore, the developed courses are needed to be evaluated according to their quality. Accreditation mechanisms, quality check, usability reviews are needed to be determined to develop effective learning environments. The finding of the case showed that the course development is not a step-by-step process. In the case, chapters in each course were developed as increments and this prevents the late finding out of errors due to analysis of the courses.

Outcomes of the SBS Case

Several principles were also extracted and developed from the findings of this case. Some of these principles are also similar or complementary with the findings of the first case like providing an effective project management team and a good project manager. Project manager’s ability to act as a facilitator, accelerator, motivator or mediator is crucial for the success of the Project. Providing a project plan which is continuously updateable is one of the most emphasized features stated by the interviewees of this case. Since it is not possible to predict everything at the beginning of the project, the need to continuous update of the plan throughout the project is needed to be realized and accepted by all team members. Related to this issue, the requirement of configuration and change management mechanisms was emphasized in this case. Other common principles with the previous case were formation of adequate teams, training for team members, quality control process and having an iterative development model. Another distinguished principle mentioned in this case was to have an effective communication infrastructure and mechanisms. Especially in geographically distributed development environments, communication infrastructure which incorporates different communication tools is needed in order to track communication data. Apart from the infrastructure, other mechanisms that define the type and way of communication among all the team members are essential for effectiveness. Rules of engagement should be determined at the beginning of the project.

Outcomes of the EPPICC Case

Based on the results of interviews conducted with the project stakeholders and the evaluation matrices applied in this case some other principles were revealed. Enabling appropriate division of responsibilities in teams is one of them. Moreover, the need for collaborative decision making in teams for easy negotiation among the members as well as enabling awareness of what is going on at any time of the project was emphasized by interviewees. Some distinguishing principles gathered from this case were careful consideration of recruitment and retention, providing a risk plan and providing usability tests. Drawing attention of students to the developed online courses or programs and keeping their attention continuous to them is an important task that should also be considered for the continuity. Planning of publication is crucial. Risk planning makes the project ready to the unexpected events. Since there was no risk planning in this project, schedule slippage occurred due to several reasons. Usability testing will enable the students focus on the content rather that the system. This also helps their retention on the courses without frustration. Trainings and iterative development approach were also emphasized as the outcomes of this case.
Outcomes of the AVC-3 Case

Supportive, directive and concerning project management was also highlighted especially for the motivation of the team members in this case. Providing a project plan which is updateable since the prediction of the later events clearly is not possible at the beginning of the project was also emphasized. In this project this could not be provided so the slippage in the schedule was occurred. People in the development teams may not have the same level of knowledge or experience about online course development as in this case. Therefore, providing training was also stressed in this case. A style guideline determination was emphasized in this case since they provide the unity of the courses that will be served in the same curriculum. This will also enable the assessment of quality of the courses easily. Conducting regular review meetings was considered essential by the interviewees as they believe that they enable the control of the project flow as well as increase the quality of materials.

The investigation of all these cases revealed some essential principles that should be included in a development model. These principles especially could be grouped in management or integration layer since they were dealt with management, planning or strategic decision making level activities. Table 2 summarizes all the principles to be included in the model.

<table>
<thead>
<tr>
<th>Management Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project management</td>
</tr>
<tr>
<td>Project Managers/Team Leaders</td>
</tr>
<tr>
<td>Continuous update of project plan</td>
</tr>
<tr>
<td>Needs assessment/ requirements analysis</td>
</tr>
<tr>
<td>Partner participation to strategic decision making</td>
</tr>
<tr>
<td>Regular team meetings</td>
</tr>
<tr>
<td>Communication/coordination mechanism/infrastructure</td>
</tr>
<tr>
<td>Conflict resolution mechanism</td>
</tr>
<tr>
<td>Accreditation</td>
</tr>
<tr>
<td>Adequate number of personnel</td>
</tr>
<tr>
<td>Risk planning</td>
</tr>
<tr>
<td>Technical support</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Integration Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trainings</td>
</tr>
<tr>
<td>Style guideline</td>
</tr>
<tr>
<td>LMS</td>
</tr>
<tr>
<td>Quality criteria</td>
</tr>
<tr>
<td>Recruitment and retention</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Micro Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogical aspects</td>
</tr>
<tr>
<td>Internal reviews as well as external reviews</td>
</tr>
<tr>
<td>Prototyping</td>
</tr>
<tr>
<td>Usability testing</td>
</tr>
<tr>
<td>Revision of the project after courses are given</td>
</tr>
<tr>
<td>Reusability of course materials</td>
</tr>
<tr>
<td>Industry standards</td>
</tr>
</tbody>
</table>

Courseware Examination

In addition to the interviews, the developed course materials were also examined. The courses developed in AVC-1 case were investigated and some slight differences in format are observed between different developers’ courses in the same team. The navigation structure of each course, the use of concept maps and the evaluation methods differ. Courses developed by the other countries involved in the project were also investigated. Some very significant differences were seen among the countries in the presentation of the content as well as the general structure.

For the EPPICC case there were no format differences among the modules since a formal structure was strictly imposed. All cases examined above were textual material with the addition of interactive examples, exercises etc. For the AVC-3 case, the structure of the courses was very different than each other, since video based lecturing was added. The presentation style of the videos was different. However, the major navigation structure was similar in all cases.
As a summary, the examination of material developed revealed that there are two basic approaches:

- Text based; with changing styles and formats
- Video based, again with changing styles.

These variations cause decrease in the quality of the courses in addition to the difficulties in assessment of the quality. Therefore, there is a need for common style guidelines which leaves some room to the developers based on the needs of their courses.

**The DONC$^2$ Development Methodology**

After the investigation of all cases as well as the literature, the final version of the model was formed. A development model based on adaptive development approach, which emphasizes iteration, concurrency, continuous feedback and collaboration was constructed. The necessary processes considered as components to develop courseware were determined. These components are positioned into three layers. Within the model, as emphasized by the circular structure used, there is no pre-determined sequence among the components, as depicted in Figure 6.

![Figure 6. The DONC$^2$ Development Methodology](image)

Iterations in the methodology are time-boxed phases, which generally take four to six weeks determined by project management, based on the amount of work assigned. All these components are assigned to different phases. This assignment does not emphasize any order as in traditional models. Any component can be assigned to a phase and can continue to be conducted in the following phases depending on the resources available. Management layer components (especially the planning activities) are conducted with a heavy workload at the very beginning of the project. Then integration layer components and individual courseware development activities in the micro layer start. All the components can be continued throughout the whole development process in different workloads. The approximate workloads of the components in phases can be seen in Figure 7. Regions with heavier loads are shown with wider black areas.

![Figure 7. Workload change of three layers during the entire project](image)
The complete workflow can be summarized in the form of phases and components as can be seen in Appendix A. At the beginning of each phase there is a phase planning meeting and at the end there are also review meetings for the evaluation and revision of what was done in that phase.

**Management Layer Processes**

The management layer deals with the strategic decision making activities of the project management team. Within the framework of management layer components, there is no pre-determined sequence. They all can be conducted concurrently as well as iteratively throughout the entire project. They can be divided to approximately four phases but this can be changed depending on the resources. These components involve tasks that are done throughout the whole development effort and they continue with different workloads till the end, as can be seen in Figure 8.

![Figure 8. Workload change of management layer during the entire project](image)

- **Project Management** is one of the main activities of the global layer which continues heavily until the end of the project and involves two major sub-activities:
  - Plan development activity is iterated several times throughout the whole process. It is revised based on the outputs of the other processes of the global layer. It does not make rigid projections for the later phases’ components but rather gives outlines for the future tasks. Therefore planning is done at the end of each phase for the following phases in the form of phase planning meeting. This planning task also gathers feedback from the continuous revision and evaluation component which will be explained later.
  - Management starts from the beginning of the project and ends only when the project is finalized. It involves the control of the execution of the plans for the phases.

- **Budget/Resource Allocation** includes two major activities:
  - Budget/Resource Planning involves the determination of the resources in terms of human, equipment and materials and their quantities for the execution of the processes.
  - Organizational Planning mainly deals with human resource planning. This activity can be repeated during the project for several times according to the availability and workload of the people when performing macro or micro layer processes.

- **Determination of the curriculum** starts with the needs assessment activity. The workload in this component is heavier at the beginning while it is less at the later stages of the project but continues until the end as small modifications may be needed for the scope of the program.
• **Coordination** is done by the project management team to provide collaboration and communication among all the teams as well as team members.
  - Communication Planning: This involves determining the information and communications needs of the people involved in the project as well as information distribution mechanisms, which enable the required information available to people in a timely manner.
  - Communication Moderating is done by a coordinator or facilitator who is also the member of project management team.

• **Configuration/Change Management** is a component which is based on the emphasis on the change in the model. Continuous communication and feedback mechanisms exist among all layers as well as all components. This requires continuous revision and change for processes.

• **Quality control** is a required component as the quality of any degree program. Project management team determine the quality criteria as well as accreditation criteria for the degree program as well as the courses in the program.
  - Quality Planning involves determination of quality policy, standards and regulations to be followed during the project and preparation of quality checklists for the evaluation of the developed curriculum and courseware.
  - Quality Inspections/Reviews is another continuous activity. This is also considered in continuous revision and evaluation component and this activity is conducted in the form of review meetings. These review meetings are conducted to check the quality of the components. In addition these will be in the form of usability evaluation meetings when they are especially conducted during the implementation of micro layer components.

• **Risk Management** is another essential component considered in global layer as this kind of development effort involves many risks that are to be overcome. Risks are generally based on hard or soft issues.
  - Risk Management planning involves deciding how to approach and plan the risk management activities for the project.
  - Risk Resolution: This activity is performed whenever any risk is occurred as a response to overcome the risk.

**Integration Layer Processes**

The integration layer deals with the curriculum development activities. Integration layer components can be conducted in two phases iteratively and some activities of the processes continue with different workloads during the project as can be seen in Figure 9.

![Figure 9. Workload change of integration layer during the entire project](image-url)
• *Determination of the courses* component of macro layer aims to determine the courses to be included in the curriculum of the degree program. Courses and their scopes are identified according to the results of the need analysis activity based on the defined structure of the degree program.

• *Decision on LMS*, that will be used to offer the courses, is an important activity. The LMS that is easy to use and handled by the course developers as well as that have many interactive features for the students is to be selected. LMS is chosen after matching the requirements with the features of the investigated LMSs.

• *Recruitment/Retention* component is considered in macro layer as it is related with the target audience of the curriculum. It is about drawing attention of the target audience as well as keeping them on site. Consideration of this is necessary if the curriculum will be served for private education or trainings.

• *Style guidelines* provide standard structure in all courses according to the quality issues determined in the global layer. As these courses serve a degree program, commonalities are required in their style.

• *Training* is required for all people who work in any of these processes. Training programs are essential to give a common insight to distributed development teams on the requirements as well as the development strategies of online materials.

• *Technology support* deals with providing continuous support for development teams in their technological issues in a timely manner.

**Micro Layer Processes**

The micro layer deals with the courseware development activities. Micro layer components are carried out continuously until all the courses are developed as can be seen in Figure 10.

![Figure 10. Workload change of micro layer during the entire project](image)

- *Needs Assessment* component involves trying to gather what changes in students’ knowledge are required by the developed course by an instructional designer.

- *Task Analysis* determines the content and the tasks necessary for the course. Subject experts are the primary source for this activity.

- *Learner Analysis* involves considering the target group’s general characteristics, their prior knowledge as well as their motivation level and attitudes.
- **Determination of Goals/Objectives** enables the identification of the main aim of the course. They define what the learner would know or perform at the end of the instruction.
- **Instructional Activities** component involves determination of activities to be included in the learning environment to provide the interaction of the learners with the material, instructor and each other based on the determined goals and objectives.
- **Content Sequencing** involves the combination of the results of the task analysis, the goals and objectives of the instruction and instructional activities to decide on the sequence of the content of instruction.
- **Evaluation Procedures** that are going to be applied for the instruction are determined based on the determined goals and objectives.
- **Searching from learning objects** component deals with searching for suitable materials in existing learning object repositories before starting to develop new materials to reduce redundant efforts.
- **Paper prototypes (storyboards)** are the paper prototypes of learning materials. This will provide to evaluate the learning objects formatively by developing about 10% of the course first to detect and reduce errors and developing the rest incrementally by supporting continuous evaluation.
- **Software prototypes (learning objects)** are the software prototypes of learning objects. All learning materials are developed as learning objects according to a determined standard. This will enable them to be re-used whenever necessary and reduce the redundant efforts. First 10% of the paper prototypes are implemented as software prototypes than the rest is developed iteratively and incrementally.
- **Integration** involves the incorporation of developed learning objects according to the determined content sequence incrementally. They form a complete course in the degree program. All courses are also integrated and form a complete degree program

**Communication**

Communication is the essential and necessary process required for all level components and for all team members. Mechanisms are to be provided to enhance effective and efficient communication to ensure timely and appropriate generation, collection, dissemination, and storage of the project information especially for the geographically dispersed team members. Collaborative services that could be included in a communication mechanism can be listed as bulletin boards, discussion boards, e-mail, e-mail notifications, online paging/messaging, chat, white board, audio/video conferencing, task lists, contact management, screen sharing, surveys/polling, meeting minutes/records, meeting scheduling tools, presentation capability, project management, file and document sharing, document management, synchronous work on files/documents (Bafoutsou & Mentzas, 2002). These services are to be provided in a “shared workspace” (Poltrock & Engelbeck, 1997) environment in order to provide the critical links among people, ideas, and information necessary for success. This workspace functions as a shared memory as in the blackboard systems used in artificial intelligence (Corkill, 1991) which aim to find a solution to a common problem by different specialists. This way, each party can post a solution and apply their own expertise to any part of the problem and contribute to the overall solution. This can be provided through a web-based shared workspace infrastructure called communication blackboard. The infrastructure will contain specific areas in each layer of the model for different teams and team members can access several of these areas according to their responsibilities to contribute to others while solving problems. The communication process is continuously supported by the communication moderating activity of the coordination process of the global layer. A coordinator is responsible for the implementation of this component.

**Evaluation/Revision**

Continuous evaluation and revision are also essential elements and conducted at all layers. The processes at all layers are continuously tested and evaluated and revisions take place as a result of these evaluations. Evaluations can be either formative or summative. As a part of formative evaluation a series of usability tests are conducted. Revisions are conducted as peer reviews or expert reviews in any of the processes. This process is supported by the quality inspections/reviews activity of the quality control component of the global layer and takes place in the form of review meetings at the end of each phase. Based on the results feedback is given to the related components when necessary.
Personnel Requirements

Applying the DONC\textsuperscript{2} methodology to an online curriculum and courseware development effort requires many roles which correspond to a different expertise. Subject Matter Expert (SME) is the specialist on the content of the course.

- Project Manager (PM) is responsible for the entire implementation of the project and who has necessary managerial skills such as leading, communicating, negotiating.
- Team Managers (TM) are responsible for the work of each team at each level activity. They also should have the similar skills with project manager and they should serve and help to the PM for the implementation.
- Coordinator/Facilitator (C) is responsible for the communication and enabling collaboration among all the team members. Whenever required, s/he should have the right to enforce even the project managers to enable the effective communication skills.
- Pedagogical Expert (PE) is responsible for the selection of pedagogical approaches that are going to be implemented in the courseware.
- Technical Expert (TE) is responsible for the technical support of the development teams.
- Subject Matter Expert (SME) the specialist on the content of the course.
- Visual Designer (VD) is mainly responsible for the style guideline defined in the project as well as the usability issues of the user-interface.
- Multimedia Designer (MD) works in the design and development of multimedia materials used in the courseware.
- Audio/Video Director (AVD) is responsible for the creation as well as the preparation of video elements that will be used in the courseware especially in the video based lessons.
- Software Programmer (SP) is responsible for the development of software versions of the courseware in the form of learning objects with the integration of them as a complete courseware.

These roles can be matched with the components of the model and this can be seen in Appendix B. Marked intersections show the role which is primarily responsible in those components in the matrix. Some additional responsibilities can be given to these roles or they can act as feedback agents in other components.

Conclusion

Virtual learning environments (VLE), in other words, online learning environments, require features different from the traditional learning environments. This issue reveals the need for careful considerations in their design and development process as the online students deserve at least the same level of quality with their traditional counterparts. A distributed online curriculum and courseware development model (DONC\textsuperscript{2}) was developed in this study to fulfill the aforementioned needs.

DONC\textsuperscript{2} was developed by the investigation of related disciplines as well as the investigation of online courseware development projects. The related disciplines that some of the principles of the model were gathered were instructional design, software engineering, more specifically the adaptive software development (ASD) model. Four different online course material development cases were examined. The model was first formulated after the investigation of the first case. Then it was applied as an evaluation framework in one of the cases and it was applied as a development framework in another case. Based on the findings gathered from these cases, the essential principles for the model were revealed and integrated into the model.

As DONC\textsuperscript{2} investigates the development effort components in different layers, they can all be adapted to the needs of the organization. It includes the important components that should be considered for the success of the development effort. It is a flexible model that does not attempt to impose rigid prescriptions to developers. It emphasizes short time-boxed iterations, collaboration at all levels among stakeholders, continuous evaluation and revision. Collaboration is essential for success and it can be achieved by effective communication which can be ensured by a coordinator who has effective communication and facilitating skills. Management based on leadership rather than an authoritative mode is emphasized for effective collaboration. Change is considered an indispensable feature so continuous updates are applied to plan.
The Model, while being improved, was applied in the final case (AVC-3) with success. Although it was not possible to put into practice all layers and components to the project, some of the components, especially, the supportive project management conducted with the help of coordination component, trainings provided to the developers and regular reviews conducted for the evaluation of the courseware were applied. The interviews conducted with the project members revealed that these components were considered as effective and helpful to the success of the development effort and quality of the work products.

As a future work, the whole model can be implemented fully to a complete program development case to investigate further if all determined components are working effectively. There is also a need to support and maintain the developed courseware to keep their quality. Therefore new courseware development cases can also be investigated to determine additional components related with the support and maintenance of the developed courseware or curriculum.

Acknowledgements

This research was supported by Human Resource Development Program of Turkish State Planning Agency

References


Appendix A

Figure A1. Overview of the methodology workflow
### Figure B1. Roles and Responsibilities matrix

<table>
<thead>
<tr>
<th>Management Layer</th>
<th>Process/Role</th>
<th>PM</th>
<th>TM</th>
<th>C</th>
<th>TE</th>
<th>PE</th>
<th>SME</th>
<th>VD</th>
<th>MD</th>
<th>AVD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Project Plan Development</td>
<td>1.1.1 Plan Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 Project Plan Management</td>
<td>1.3 Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Budget Resource Allocation</td>
<td>2.1 Budget Resource Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2 Determination of resources Cost estimation Cost budgeting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3 Determination of the program Needs analysis for the program</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4 Determination of Quality Criteria</td>
<td>5.1 Determination of Quality Criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2 Determination of Quality Issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1 Determination of Needs</td>
<td>6.5 Needs assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2 Determination of Needs analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.3 Continuous training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.4 Examination of Style Guidelines</td>
<td>12.1 Examination of Style Guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.2 Determined guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.3 Technology support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.4 Internal support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.1 Determination of responsibilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.2 Technical support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.3 Technical support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Needs Assessment</td>
<td>14.1 Needs Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Learner Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Determination of course objectives</td>
<td>17.1 Determination of course objectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Determination of instructional activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Determination of Evaluation Procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Sequencing the content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 Searching from LOs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 Paper prototype preparation of 10% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Software prototype preparation of 10% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 Integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Searching from LOs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 Paper prototype preparation of 30% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Software prototype preparation of 30% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 Integration of 30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 Searching from LOs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 Paper prototype preparation of 60% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 Software prototype preparation of 60% of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32 Integration of 60%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 Integration of Courses for the Program</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End of course development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 Integration of Courses for the Program</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End of program development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Asynchronous discussion forums: success factors, outcomes, assessments, and limitations

Martin A. Andresen
School of Criminology, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6 Canada // Tel: 778 782 7628 // Fax: 778 782 4140 // andresen@sfu.ca

ABSTRACT
Online learning has been burgeoning over the past decade with one of the more popular modes of conducting online learning being the asynchronous online courses. Within the asynchronous online course, the asynchronous discussion forum replaces the face-to-face interaction of the traditional classroom, but is this form of discussion able to enhance the learning process? This paper reviews the literature regarding asynchronous discussion forums finding that the asynchronous discussion forum is able to generate the critical dimensions of learning found in the traditional classroom, but it has its limitations.

Keywords
Asynchronous discussion forums, On-line learning

Introduction
The process of discussions is a critical dimension of the learning process. Moreover, the learning experience itself has been shown to be enhanced through the regular participation in discussions (Kolb, 1984). Whether these discussions take place in a traditional classroom or through online teaching using electronic means, their importance is integral to both learner achievement and learner satisfaction (Fulford and Zhang, 1993; Zhang and Fulford, 1994). Because of the development of faster computers, improved telecommunications networks, and the development of readily accessible software the availability of courses and programs through online teaching has been growing exponentially.

Regarding the growth of offerings in online education, Tucker (1995) found that the percentage of colleges and universities in the United States offering online education went from 3 – 30 percent, 1990 – 1995; and Gubernick and Ebeling (1997) found that the number of institutions in the United States offering online education increased from 93 to 800, 1993 – 1997. No doubt, in more recent years, the growth continues to be exponential, not only in the scale, in terms of how many colleges and universities are offering some form of online education, but in the scope of courses and programs that are available in each of those institutions.

In a more recent study funded by the National Center for Education Statistics that investigated the 2000 – 2001 academic year, Waits and Lewis (2003) found that 90 percent of public post-secondary institutions in the United States offered distance education, with 90 percent of those institutions undertaking asynchronous online courses. These numbers led to more than 3 million learners (82 percent undergraduates) being enrolled in almost 120 000 credit-granting courses (76 percent undergraduate) that year—over 127 000 courses if one considered non-credit courses (Waits and Lewis, 2003). Additionally, and important for this review, of those institutions that would offer or planned to offer distance education within the next three years, 80 percent stated that they would increase or start using asynchronous online education as the primary mode of dealing with those courses. Though no study was found that measured online education in Canada, similar percentages are expected. As such, understanding the nature of asynchronous online education is critical because of its widespread use and expected expansion; in particular, it is important to understand the determinants of effective learning in an asynchronous online discussion because these discussions are the equivalent to the face-to-face discussions common in the traditional classroom that Kolb (1984) found to be critical in the learning process.

There are some obstacles to overcome that are specific to an asynchronous online discussion and, hence, its learning process. All distance education, whether online or not, is defined by having the instructor and learner separated in space (Mood, 1995); with the added dimension of an asynchronous discussion, they are also separated by time (Carswell and Venkatesh, 2002)—of course, it is possible that multiple learners and the instructor may be online at the same time enabling an asynchronous discussion to occur very close to “real time”. Consequently, the asynchronous discussion forum must be specifically analyzed in order to enable the asynchronous discussion to be as
(or more) effective as the traditional face-to-face-discussion if high levels of learning are to take place—some of the benefits of the asynchronous discussion that may make it more effective than the traditional face-to-face discussion are that it allows those people who need more time to participate to contribute to a discussion, a discussion participant cannot be “cut off”, and there is a transcript of the discussion for study purposes after the discussion takes place.

Despite this need for analysis of the asynchronous discussion forum, the literature is growing, but relatively small and spread across a wide array of disciplines ranging from education to physics to philosophy. In this literature review, I cover the research areas that I consider the most important: the makings of a successful asynchronous discussion, assessing asynchronous discussion forums, and the limitations of asynchronous teaching. These sections are then followed by conclusions and directions for future research.

The Components of a Successful Asynchronous Discussion

Making a successful asynchronous discussion is probably the most important aspect for an instructor to consider. Though assessment and the limitations of these discussions are also important (see below), if one is truly concerned about the generation of knowledge in learners the asynchronous discussion itself is key. Two components in the literature emerged as being particularly important for a successful asynchronous discussion forum: the role of the instructor and achieving deeper/higher learning. Each is discussed in turn.

The Role of the Instructor

Instructing in an online environment is inherently different from the conventional classroom. Therefore, we should expect that the roles that we take on as instructors will change, to some degree, when we deliver an online course and monitor an asynchronous versus a live discussion. Coppola et al. (2002) investigate this change through interviews of twenty faculty and found that instructor roles in cognitive, affective, and managerial activities changed in the online environment.

First, the relationship between the instructor and the learner (affective role) needed to change because of the loss of face-to-face contact in the conventional classroom. Instructors needed to find new ways to express emotion, or passion for the subject matter, when communicating ideas to the learners. As a consequence of this search for new modes of communication, the instructors felt that the virtual classroom, including the asynchronous discussion forum, became more intimate. Second, as a result of instructors searching for new modes of communication, instructors found that their teaching involved deeper cognitive complexity (cognitive role). Tone of voice, body language, and spontaneous questions to clarify concepts are all lost in an asynchronous learning environment necessitating the instructor to become much more cognitively involved in the learning materials. And lastly, the managerial role changed through differences in class and course management. Again because of that loss of expressiveness and spontaneous questions in the face-to-face classroom, instructors need to pay more attention to the details within the course: more precision and formality in setting assignment expectations, for example. Clearly, knowing that successful teaching in an online environment necessarily involves changing your role as an instructor in multiple dimensions is important information, but other research has found specific details of what instructors need to do in order to have that success.

Simply forming an asynchronous discussion forum, providing the technology, and a question or topic of discussion is not enough to ensure success in an asynchronous discussion (Guldberg and Pilkington, 2007). Though there are always factors beyond the control of the instructor such as the personalities of the learners enrolled in the asynchronous discussion forum and how they chose to interact with other learners (Guldberg and Pilkington, 2006), there are two questions that emerge and need to be considered: first, and related to the managerial role referred to by Coppola et al. (2002), what does the instructor need to do in order to stimulate good asynchronous discussions and, second, once that discussion is underway how much should the instructor intervene?

Factors that are within the control of the instructor and have positive effects with the generation of complex discussion are the time learners have to prepare for a discussion, the time needed to develop online relationships within the asynchronous discussion forum, and the nature of the discussion (McConnell, 1994; Salmon, 2002). More
specifically, Dysthe (2002) has found that learners should have a reading assigned to them, have time to reflect on that reading as well as a discussion topic or question, and then present examples (real or hypothetical) that relate to the topic or question to the other learners and defend those examples in the discussion. Moreover, successful questions or discussion topics must be related to the learning objectives with clarity in due dates, expectations, and the weighting of grades so that learning objectives may become learning outcomes (Guldberg and Pilkington, 2007; Majeski and Stover, 2007). And as found by Fung (2004), when discussion questions or topics were specific and related to a concept or idea within the course readings the discussions were more successful in generating complex interaction between learners than those discussions that were begun with open-ended and broad questions. For example, asking a learner what to do in a situation rather than what they thought of a situation generated complex interaction.

With regard to the timing of the discussions, time to prepare for a particular discussion is only one of the important temporal considerations needed to be considered by the instructor. Rather than beginning the course discussions with deep cognitive questions and topics, questions and topics that encourage discussions on social, personal, and reflective levels should be used in the beginning discussions in a course. This strategy is used to initiate discussion between learners in order to develop online relationships. Only once these online relationships have been developed should cognitive questions and topics be gradually introduced to the learners (Salmon, 2000).

Equally important to the design of the asynchronous discussion forums is the level of intervention taken on by the instructor. The research on this dimension of the asynchronous discussion forum is quit clear for the instructor when considering the learning outcomes of the learners: back off. In a study that asks the question of what role an instructor should undertake in an asynchronous discussion forum (sage, guide, or ghost), Mazzolini and Maddison (2003) found that it depends on what the instructor wishes to accomplish. Learner ratings of a course will show that an instructor is more enthusiastic and expert if s/he increases his/her postings. Similarly, Swan and Shih (2005) find that the perceived presence of an instructor is more important than the perceived presence of peers in student satisfaction. However, an instructor that contributes significantly to a discussion tends to decrease the length of discussions (this does not necessarily decrease the quality of the discussion, however) as well as their frequency. What appears to be occurring in this situation is that the instructor can decrease learner – learner interaction because the learners begin to rely on the instructor to answer questions, becoming the export or sage to “settle” debates (Guldberg and Pilkington, 2007; Paloff and Pratt, 2001).

Rather, the instructor should intervene, but only in order to keep the discussion on track, or take on a cheerleading role to motivate the discussion (Dysthe, 2002; Paloff and Pratt, 1999). The role of a cheerleader or motivator is critical because it is the learner-learner interaction that truly engages with ideas and relates back to Kolb’s (1984) statements that discussion is a critical component of the learning process, not waiting for the answers to fall from Heaven’s academic prophet. Preferably, the instructor should spend his/her time preparing materials and the carefully thought out discussion questions and topics that relate to learning objectives, as discussed above.

**Achieving Deeper/High Learning**

The ultimate goal of spending the time to develop an asynchronous discussion forum, and manage it in the appropriate manner, is to create an online learning environment that will achieve high levels of learning. One way to assess whether or not this has occurred is to test the level of learning that has been reaching within the discussion using an appropriate methodology.

The earliest study found that investigates the level of learning is Webb et al. (2004). This study is not concerned specifically with the dialogue (quantity and quality) that occurs within the asynchronous discussion forum, but the learning outcomes measured by learners’ grades for the course. Webb et al. (2004) finds that as participation in the asynchronous discussion forums (measured by access to the discussion forum and the number of postings) increases so do the measured grades for the learners. The limitation with this finding is that it does not consider the degree of cognitive engagement in the asynchronous discussion forums. If a higher degree of cognitive engagement is achieved then perhaps all learners will extract a greater benefit from the asynchronous discussion.

Using a number of different models of hierarchical learning, Schellens and Valcke (2005; 2006) measured the degree to which asynchronous discussion forums reached the higher levels of knowledge creation. They found that...
asynchronous (versus synchronous) discussion forums attained a higher proportion of higher phases of knowledge creation. This occurred because the vast majority of communication in the asynchronous environment was task-oriented, greater than 88 percent (Schellens and Valcke, 2006). Additionally, they found that groups with more discussion resulted in higher level of knowledge construction (Schellens and Valcke, 2005), as well as smaller asynchronous discussion groups (n ≤ 14) resulted in higher levels of task-oriented communication and, hence, a higher proportion of higher phases of knowledge creation (Schellens and Valcke, 2006). This last finding is important for the original development of courses and their corresponding asynchronous discussion forums. Though some instructors may fear their job security is at stake with the advent of online education, this research clearly shows there are limitations that should be put in place on class and discussion sizes for reasons of pedagogy.

Lastly, and related to the previous subsection, Zhu (2006) has found that high levels of interconnectedness between learners leading to higher levels of knowledge construction must be explicitly built into the discussion assignment and nurtured by the instructor. More specifically, the instructor’s discussion design is more important that any specific technology used for the asynchronous discussion forum. Knowledge construction only occurs because of careful planning: clear, well-defined, well-crafted questions and discussion topics. Without such planning and subsequent guidance, only lower levels of cognitive engagement will occur. As noted by Howell-Richardson and Mellar (1996), the level of interaction between learner and, hence, their cognitive engagement, may be increased with only slight modifications to the course design and the instructor’s behavior.

Assessing Asynchronous Discussion Forums

Given the importance of the discussion in the learning process at both the theoretical and empirical level, an appropriate measure of participation should be a component of each learner’s grade for the course. Though there are always learners who wish to participate in discussions, face-to-face or online, many learners need an incentive to participate in class discussions. As such, just as it is critical for the instructor to set out each task in an asynchronous discussion forum clearly and succinctly (see above) so must be the assessment of those asynchronous discussion forums in order to facilitate as much discussion as possible between the learners.

The primary difficulty in making any assessment of an asynchronous discussion forum is the huge volume of data that are available to be assessed—content analysis has been shown to be useful in a small class setting (Bali and Ramadan, 2007). Within this vast amount of data, learners decide where to post their comments, making the discussion not follow the temporal sequence of the postings—for example, if there are five postings within one discussion forum, a learner may respond to the second posting after having read all five of the postings (Dringus and Ellis, 2005). Consequently, discussions appear to the reader as fragmented and discontinuous because of this temporal separation of postings making assessment complicated. As a result, there is no one accepted method for assessing learner participation in asynchronous discussion forums.

However, this is not to say that there are no options. Roblyer and Wiencke (2003) provide an assessment rubric for such a learning environment, but this rubric is far too complex to analyze even a small number of long asynchronous discussion forums. Alternatively, and particularly relevant for large number of asynchronous discussion forums, the frequency of logging on to the online environment and the length of time spent in the online environment may be used as an assessment tool that is readily and easily available within many of the online systems (see, for example, Ahern and Durrington, 1996; Taraban et al., 1999). However, the quality of time spent in the online environment is a just as important a consideration in order to assess participation in asynchronous discussion forums (Dringus and Ellis, 2005)—one may simply log in multiple times a day while playing computer games and chatting with friends online. It is very possible for a learner to have a great influence on the quality of a discussion from short participation times within the online learning environment; rather than spending all of his/her time staring at the computer screen, s/he may read the discussion, log off, think of the issue, and then quickly log on to post a response. As a result, the whole (the actual contribution of an individual learner) may in fact be greater than the sum of its parts (the apparently short log in times and postings).

The challenge, then, is to measure the quality of a learner’s contribution to an asynchronous discussion forum. Dringus and Ellis (2005) suggest that instructors need to know the following: when a learner posts in the asynchronous discussion forum relative to the assignment time frame, whether or not these postings are responses, if they are responses were they immediate, and whether or not a learner’s postings generate responses from other
learners. Gathering this type of information, however, is difficult and time consuming if performed manually. Herein lays the need to generate methods of assessment for asynchronous discussion forums that can apply more complex assessment rubrics (see Roblyer and Wiencke (2003) from within the online system itself rather than relying on manual interpretation of the discussions (Garrison et al., 2001; Jarvela and Hakkinen, 2003)).

One method of assessment for asynchronous discussions that has been proposed recently is the use of data mining. Data mining literally analyses large volume databases in order to extract any relationships, clusters, and/or patterns to the data (Dringus and Ellis, 2005; Romero and Ventura, 2007). Even simple data mining procedures can generate valuable information regarding such factors as: the number of learner postings within a particular time span, whether or not a learner initiates discussions and/or responds, and how long it takes a learner to respond to an initial posting (Dringus and Ellis, 2005). This information may be used as the entire assessment of a learner’s participation in the asynchronous discussion forum or as a component of that assessment; alternatively, this information may be used to see if it is worth the time to further analyze actual transcripts of a learner’s participation.

The difficulty with employing data mining methods in asynchronous discussion forum evaluation (or any evaluation for that matter) is that currently data mining tools are too complex for those not trained in computer science (Romero and Ventura, 2007). Consequently, there is a need for collaborative efforts between computer scientists, educators (which includes computer scientists, of course), and the providers of the technology for online education in order to develop data mining tools within the educational technology that can be used by the average user with minimal, or no, training. This relationship can also become symbiotic and dynamic in the sense that an initial assessment rubric can be developed with the current constraints of data mining in the asynchronous discussion forum context. In time, armed with a “wish-list” from those teaching in an asynchronous environment, computer scientists can develop more/better data mining tools to be implemented into the online teaching technology.

The Limitations of Asynchronous Teaching?

Despite the fact that there are clear methods for establishing, moderating, and maintaining successful asynchronous discussion forums have been outlined above, as well as a move toward methods of assessing those asynchronous discussion forums in a more meaningful way, one may still ask whether there are limits to the use of asynchronous discussion forums in education. Surely no one will argue that the lecture hall has limits: it is not conducive to class discussions; and similarly, the tutorial classroom (with only desks, tables, and chairs) is not a suitable environment for the lab component of a chemistry class. Then what limitations, if any, are there for the asynchronous discussion forum?

There appears to be very little research on the limits of asynchronous discussion forums, but the research that is available is consistent, at least when it comes to the type of discussions that are feasible in an asynchronous learning environment. Though in general the sciences have issues regarding the feasibility of asynchronous discussion forums (Larreamendy-Joerns and Leinhardt, 2006), two studies have emerged that are consistent in showing that problem-based learning is difficult in an asynchronous discussion forum. Kortemeyer (2006) found in an introductory physics course that conceptual issues relating to physics education work well within an asynchronous discussion forum: discussing new terms, concepts, etc. However, when it came actual problem solving (usually mathematics), critical in much of the sciences, the asynchronous discussion forum was not successful—a similar result was found by Hong et al. (2003) regarding problem solving in an asynchronous discussion forum in a statistics class.

Though some learners that complain about participating in asynchronous discussion forums simply prefer to work alone rather than in groups (see Oliver and Omari, 2001), there are reasons why problem-solving discussions are difficult in an asynchronous environment. The discussion of concepts and ideas, though needing to be focused, are able to waver, whereas discussions regarding problem-solving tend to be extremely specific: what am I missing in order to solve this equation? In other words, posting a question or response regarding a concept, and waiting for someone to reply, is much different than waiting for someone to reply to your call for aid in solving a problem—one has a very specific desired end and the other does not, necessarily. Consequently, instructors need to be aware not only of the specifics, deadlines, and weighting of an asynchronous discussion question or topic, but whether or not that type of topic or question is appropriate in an asynchronous environment.
In a completely different type of course, multicultural education, Merryfield (2001) came across a paradox in her asynchronous discussion forums. On the one hand, the participants in an asynchronous discussion were found to be far more reflective, frank, and willing to discuss sensitive multicultural issues regarding topics such as racism, white privilege, and homophobia than similar discussions in her face-to-face version of the same course. Merryfield (2001) believed this to be because the learners did not have to look the other people in the eye, which allowed them to speak more freely. But despite the deeper and more engaging asynchronous discussions, as opposed to Merryfield’s (2001) traditional face-to-face classroom experience, most of the learners in this course felt that the asynchronous discussion forum was a less meaningful form of communication. Learners felt disconnected from the discussions and were left wondering if the experience was actually real.

This is indeed a strange finding because, as discussed above, more engaged discussions generally do not emerge until online relationships are established (Salmon, 2000). But if these online relationships were established in Merryfield’s (2001) course in order to have these more engaged discussions, what then is the definition of an online relationship? As noted by Hillman et al. (1994), all interactions in an online learning environment (including the asynchronous discussion forum) are mediated through technology. Consequently, a question that needs to be answered is: what is the nature of the online relationship and does it limit the scope of the use of the asynchronous discussion forum in online educational environments? Though this question cannot be answered here, Shea et al. (2005) suggests that the instructor’s role is of the most promising mechanisms to establish online learning relationships. This harks back to the earlier statement that instructors need to spend time preparing the asynchronous discussions rather than being active within them.

Conclusions and Directions for Future Research

Online education has experienced a vast expansion in the creation of its environments, including both the current and expected use of asynchronous online courses and their corresponding discussion forums. The research on the asynchronous discussion forum presented here has shown both consistent results in what needs to be done for their present success (what makes a successful asynchronous discussion forum) as well as what needs to be done for their continued success (the assessment of asynchronous discussion forums and their limitations).

It is clear that asynchronous discussion forums can achieve high levels of learning, but people in decision-making positions must be aware of the conditions for this to occur. Asynchronous online courses are not a method of displacing instructors. In fact, because of the nature of successful asynchronous discussion forums, asynchronous online courses need to be as instructor intensive (instructor to learner ratio) as the traditional classroom. And in some cases, namely problem-based learning, the asynchronous discussion forum does not appear to be appropriate. In other words, there is no one size fits all application of asynchronous online learning. Consequently, asynchronous discussion forums may form part of a more generalized model of learning—a blended learning approach.

Though some of the early research on asynchronous learning environments, including the asynchronous discussion forums, has been accused of lacking in both quantity and quality (see Wegner et al., 1999, Kyounghee Lim, 2001), this is clearly no longer the case, particularly for the quality of research. Though the research on the nature of a successful asynchronous discussion forum appears to be quite clear at this time, there are, however, avenues of research that do need to be pursued or elaborated in regard to the assessment of the asynchronous discussion forum and its limits.

As indicated above, the assessment of asynchronous discussion forums is a large undertaking when the discussions are of a significant length and/or number and there are a lot of learners involved in the discussion itself. Data mining definitely provides a fruitful method of assessment that can evolve through the collaboration of educators and computer scientists; as data mining techniques improve and respond to educators hopes for assessment, so does the ability of educators to design and implement asynchronous discussion forums that are conducive to the learning process and can also be assessed with relative ease. The difficulty is in the user-friendly nature of these tools. Clearly, these tools need to be integrated into the online teaching technology in order for them to be adopted en masse. Future research in this area needs to investigate the ability of new data mining tools to be incorporated into the technology of online education while at the same time being sensitive to availability: not just the relatively large corporations such as Blackboard and WebCT, that recently merged, but also the smaller, free, and open-source course management systems such as Moodle.
On a related note, the availability of free management systems such as Moodle may be critical in assessing benefits to costs for the practicality of delivering asynchronous online learning. However, one must be careful as this is a limited view of online learning. The tuition collected for asynchronous online learning may cover the operational costs justifying the existence of an online instructional department, but is the value of the outcomes justified by the cost? In other words, is the value of an online learning program worth the cost of delivery? This topic is another direction of future research.

With regard to the limitations of the asynchronous discussion forum, one future research direction has already been posed above: an investigation of the nature of the online relationship. However, the physics and statistics examples above also point to the need for research on setting the boundaries of online education based on pedagogical grounds—the community of inquiry model may provide considerable guidance here (Garrison et al., 2000; Garrison et al., 1999). No medium or method can be all things to all people and to all course material so the educational community needs to be aware of when and where to stop the development of asynchronous discussion forums specifically, and online education, generally.

References


Using webquests for oral communication in English as a foreign language for Tourism Studies

Jesús García Laborda
Department of Applied Linguistics, Polytechnic University of Valencia at Gandia // jgarcial@upvnet.upv.es // Tel. +34 666 28 45 57 // Fax. +34 962849309

ABSTRACT
A long-standing debate in native and foreign language learning revolves around the use of computers to promote genuine social and professional communication. Webquests are a very common way of using Web resources to research a variety of topics, and if appropriately used can trigger the situations necessary to develop both written and oral communication. A webquest is a research activity that requires the learner “to collect information about a subject using the web” (Sharma & Barrett, 2007: 24). Much has been said about synchronous and asynchronous communication but relatively little research has been carried out on the effect of webquests in in-classroom communication. The aim of webquests is to motivate and promote students’ critical thinking in order to solve problems or develop projects. This paper presents a practical background for using webquests, mainly orally. Oral use of webquests is rarely addressed in research and is especially valid for students of foreign languages for professional purposes, and specifically English. The framework hereby presented addresses three main aspects of the use of webquests: as a social constructivist learning approach, as a professional development activity and as an oral development task. If the use of webquests is associated with students’ professional needs, their implementation can be very successful and help students’ skills both in language development and cooperative work. This paper also provides a real example to facilitate the understanding of the theoretical concepts.

Keywords
Webquests, oral development, English for the professions, Tourism

Introduction

In the last few years, Schools of Tourism have increased in number everywhere in Europe from Ireland to Spain. Although five or six years ago, many expected written communication (synchronous and asynchronous communication) to prevail in specific purpose foreign language education, recent developments in technology have changed this belief dramatically. The power of images transmitted by Internet video conferencing and the possibility of attaching video files to written messages have turned teachers towards the vital need to develop students’ oral production, more specifically in English as a foreign language for tourism. Additionally, all the latest Internet-based technology requires certain training that not all students bring to the class. In a school experiment done only three years ago at the Polytechnic University of Valencia, Spain it was noticed that many students lack the appropriate necessary resources to manage themselves electronically (García Laborda, 2002). Thus, they can chat and send emails every day but they have difficulties with the use of Internet for professional development (García Laborda, 2004). Estimates of the number of students who can work using a foreign language properly after graduation in activities such as searching for tickets, contrasting tourism information or finding economic studies or reports online vary according to different teachers in many Tourism colleges. Colleges like Gandia (Universidad Politécnica de Valencia) have introduced many subjects devoted to the specific use of computers for Tourism including Amadeus or Computers I & II. Additionally, foreign language classes have started to include computer skills use for communication. an important component in new technologies for communication. To do so, asynchronous communication experiences started in 2000 in cooperation with Valdosta State University. They seemed to work efficiently for a while but introducing speech communication procedures has been more difficult. Apart from issues of security, there is the additional difficulty of controlling the students’ talk. A further problem has been the different perspectives towards the role of the Internet in professional communication which are frequently used in learning tasks in Content Based Language Learning such as contrasting technical business communication or, for instance, a meeting to negotiate bulk rates between two travel consulting companies versus the ordinary customer attention, as in a travel agency simulation.

Some of these problems of training and communication have traditionally been solved through students’ simulated interviews, role plays or other classroom tasks. However, it is not unusual that these exercises become decontextualized because the market nowadays changes easily. For instance, air travel varies greatly according to the
changing price of fuel. Therefore, it is necessary to provide students with the required information that can place them in the type of near-to-real situations that they can expect to face upon their university graduation. As a consequence, the speaking needs for most Tourism students are (García Laborda 2002): professional speech, contextualized tasks, efficient and current information and opportunities to interact with students of different nationalities. In addition, students also need a certain fluency in their expression, and the capability of using adequate vocabulary acquired both passively and actively. This vocabulary is usually acquired through oral interaction but also through reading (Shanahan, 2006; Krashen, 1998; Cho & Krashen, 1994; Krashen, 1989). The Internet offers a unique opportunity for natural unmodified input that textbooks can seldom match. Finally, the Internet has a significant role in promoting Content Based Language Learning in English classes for business (Luzón Marco, 2002, 2001).

From the Students’ Needs to Webquests: Theoretical Support

Luzón Marco (2003) has given examples of different uses of webquests in the teaching of ESP. A webquest is a research activity that requires the learner “to collect information about a subject using the web” (Sharma & Barrett, 2007: 24). However valid some of her research might be as descriptive principles, Luzón Marco has clearly failed to provide a theoretical support to the use of webquests. Even more significantly, few if any researchers have been able to connect the use of webquests with students’ oral development.

Luzón Marco (2002) states that webquests are valid “to promote technological literacy among students and the focus on content-based instruction” (20). It is generally accepted that there are many reasons to support the use of webquests in Content Based Language Learning (Figure 1) – as a social constructivist learning approach, as a professional development activity and as an oral development task.

Figure 1: A theoretical diagram of the use of webquests
As a Social Constructivist Learning Approach

In relation to behavior development, webquests facilitate a change of habits and build knowledge internally through experience and the interaction between students’ free and conditioned moves. Given the importance of the social context in which oral interaction occurs and cooperative attitudes of students negotiating meanings and reflecting on their own performance, collaborative computer use and the importance of professional and intercultural competence, the use of webquests for language development relies mainly on the theory of social constructivism as expressed by Vygotsky (1978), Bruner (1996) and Lantolf (2000), along with Bandura’s social cognitive theory (Bandura, 1977; 1986; 2001; 2005). Through the continuous use of specific dialogues (and their internalization in the students’ own knowledge) based on results from webquests, tourism students learn to manage themselves first in professional real-like communication situations (Turcotte, 2005; Lachter & Bever, 1988) and then in reality. Classmates also help through this interaction to shape the idea of the professional world and increase their knowledge in relation to the situations that the student will have to experience when at work. An example of this would be the repetition of conversations between clients or potential partner companies in meetings and debriefings, or the search for fares (Appendix 1 presents a specific case with a webquest based in a class situation based on oral development for travel agencies that will be addressed in the rest of the article). The finding and use of web sites and their management leads to controlled learning (Wennik, 2004; Seamon, 2001) and, even more importantly, to increasing experience and capability in facing similar professional situations. The fact that students have to interrelate with other students for presentations and meetings benefits social interaction and increases motivation (Kennedy, 2004; Milson & Downey, 2001; Al-Bataineh et al., 2000).

Cooperative and Project-based learning are probably the best uses of webquests in professional training. García Laborda (2003) states that both language learning methods (cooperative and project-based) are in high demand in the contemporary professional world. The benefit of using webquests in this way is that students are required to interact with others to solve problems of information gathering or resisting difficulties (Milson, 2001). For example, given that they are going to prepare a joint trip, they will have to support themselves to find places and times that can best fit other elements such as hotels, travel or even attractions. In this way, students can establish learning connections between the real situations of regular ticketing or even interviews with customers with their regular class tasks. This is because students need to get immersed in real-life situations that are seldom found in the classroom. Besides, webquests favor project and cooperative learning because when “learners work in pairs or in teams, they find they need skills to plan, organize, negotiate, make their points, and arrive at a consensus about issues such as what tasks to perform, who will be responsible for each task, and how information will be researched and presented” (Moss & Van Duzer, 1998). Additionally, researchers in foreign language instruction assert that cooperative language learning benefits less advanced students increases the sources and variety of input and helps to transfer search and language strategies (Francis; 1999; McGroarty, 1989). In the case of Tourism students, they can help one another with searches, stating the results and rehearsing the presentations. However, teachers and instructors should make a special effort when preparing webquests to ensure they choose an appealing and professional topic, make clear plans and pathing, include research tasks, emphasize oral production, and, whenever possible, try to ensure written task production as well as planning activities to share the results with the group / class.

As a Professional Development Activity

Luzón Marco also states that webquests are valid in English as Foreign Language learning when they are employed to solve a problem (2001). Indeed, good webquests should be realistic (Lamb, 2004; Boswell, 2003), following Krashen’s (1982) advice that when the tasks are meaningful, they produce much better learning. After all, when dealing with webquests, students will be working with real materials mostly distributed through the web. Thus webquests are intended to simulate real-life situations. If for example a travel agent is required to produce a specially tailored travel package, it would not be so different from the example addressed in this article. Another advantage that highlights the value of using webquests in tourism is the facility of working with real data. In 2002, García Laborda (2002) observed that tourism students usually do not have a realistic perception of professional life. They are usually ignorant of basic knowledge that may be of significant importance in their first job such as pricing, knowledge of different companies, international resources and internet search skills, while displaying limited computer skills (other than word processors and a/synchronous communication devices). All this information can be accessed immediately through the web and, more importantly, it needs to be reflected on and internalized in the students’ knowledge, attitudes and performance (Al-Bataineh et al., 2000). This also leads to a change of attitude towards the use of the
Internet (Brown & Warschauer, 2006; Dudeney, 2003). Therefore webquests are a valid means of changing from a traditional perspective on language teaching for specific languages to content-based language teaching (García Laborda, 2004; Luzón Marco, 2001).

The use of webquests in English leads directly to the teaching of regular professional contents in L2. Content-based language learning has certainly acquired a greater relevance due to the integration of information technologies in foreign language learning. In fact, this has transformed the traditional vision of content-based language teaching towards the new idea of CLIL (Content Language Integrated Learning) (Dalton-Puffer & Nikula, 2006). It is commonly accepted that English teachers usually have difficulties integrating content in their classes because they necessarily need to “fake” knowledge of the specialties they do not usually possess. In other words, it is unusual to see a medical doctor teaching L2 for students of medicine, but the Internet can sometimes help to facilitate this lack of knowledge through redirecting the classroom tasks as students have the opportunity to search for their own information about a topic. The Internet provides students with the real materials they genuinely need in order to work or understand. Webquests, apart from being useful, are also challenging, meaningful and lead to experiential learning. They can also be very amusing, and develop students’ efficiency in gathering professional information and increase their motivation (Warschauer 1996). As a result, a student or group of students who can accomplish a webquest and propose solutions will have solved the language problems to obtain a certain product (in this case, the final report and the concluding presentation) (Warschauer 1999).

From Professional Content Learning to Web-Based Language Development

If a foreign language is better learned when it is meaningful and acquired experientially as Krashen (1982) asserts, content learning should promote the acquisition of conscious behavior (see above) and language that is specific to the field. Therefore tourism students should try to familiarize themselves with the different topics related to their field, with its vocabulary and certain typical communicative acts such as meetings, negotiations and presentations. Computers provide the students with language that is sometimes accessible but is most often just above their competency level leading to what Krashen (1982) called L2 acquisition. It also provides the continuous opportunities to access the L2 that, in not a few cases, are difficult to find outside the classroom when the student does not experience the anxiety of the instructional classroom drills or exercises. This is especially true for reading skills. As a consequence, the possibility of learning tourism through the class and also to complement it with language opportunities outside the classroom increases the value and utility of foreign language learning at the same time as it prepares the students. Accordingly, students working with a railways website are likely to practice their oral skills with dates, train facilities (very similar to hotels and planes), timetables and many other practical aspects of language at the same time as they learn to recommend a trip or a segment based on the same times and prices. This combination also allows students to use professional as well as general non-specific language. Professionally speaking, it gives insights into fares, train changes, train classes, train companies, bookings, payments (use of credit cards on the web) and the types of services usually associated with railway services such as hotels and car rental, along with those services usually associated with travel companies. It also has a positive effect on learning the specific language faster. When students can integrate topic learning with language learning they tend to associate both and as a result when communicating in L2 more fluency can be expected. Additionally, the new knowledge can be very useful not only for their business and personal lives but also for other university subjects such as “Travel Markets” or more practical topics of study such as ticketing or the use of the AMADEUS reservations system.

As an Oral Development Task

There are very few articles that have shown how computers and the Internet promote oral acquisition. However, more research has tried to evidence the effects of reading on oral performance (Caravolas & Bruck, 1993; Krashen, 1989). Nobody can deny that language input favors language performance. Today, it is broadly accepted, as Krashen and others have repeatedly insisted, that the acquisition of vocabulary through reading can be reflected in students’ oral performance. After all, even in the structuralist and behaviorist approaches to language teaching, it is assumed that all types of input have an effect on oral production (Benati, 2004; Krashen, 1989).

A quite different issue is how computers can facilitate language learning and whether uncomprehended input, such as the learner sometimes reads on internet web sites, can be incorporated into the learner’s speaking repertoire.
Although it can be expected that comprehensible and uncomprehended input can facilitate oral performance as has been shown with regular writing, this paper tries to show how webquests can be used to promote interaction opportunities. The following diagram shows the process of how the webquest is approached by the students.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Effects on learning</th>
<th>Effects on oral development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webquest is presented to the students</td>
<td>Part of the language and structures are learned</td>
<td>Oral and written input is received from the teacher</td>
</tr>
<tr>
<td>Students meet and assign roles</td>
<td>Cooperative work, task assignment, social interaction, motivation</td>
<td>If some part of this work is done in class they are likely to do it in L2</td>
</tr>
<tr>
<td>Students seek information individually (but in contact)</td>
<td>Passive and active reading, structure and vocabulary learning, negotiation and support (through cooperation), professional development (getting to know the market)</td>
<td>New vocabulary is apprehended. It will probably be used in oral performance later</td>
</tr>
<tr>
<td>Students have a meeting to propose the package (better in class in front of the teacher)</td>
<td>Learners interchange information, social interaction, passive and active reading, structure and vocabulary learning, output after learning</td>
<td>Previous organization and oral rehearsal</td>
</tr>
<tr>
<td>The group produces a report / booklet</td>
<td>Learners interchange information, social interaction, passive and active reading, structure and vocabulary learning, output after learning</td>
<td>The written output will be incorporated into the final presentation</td>
</tr>
<tr>
<td>A presentation is given to the rest of the class</td>
<td>Learners interchange information, social interaction, passive and active reading, structure and vocabulary learning, output after learning</td>
<td>Previous organization and oral rehearsal</td>
</tr>
<tr>
<td>(Alternatively, students could have a debriefing with a possible customer interested in the product)</td>
<td>Learners interchange information, social interaction, passive and active reading, structure and vocabulary learning, output after learning</td>
<td>Presentation in front of the teacher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feedback session</td>
</tr>
</tbody>
</table>

**Learning Processes**

As can be observed in Figure 2, oral production and development is fostered through reading and oral interaction - both comprehensible and incomprehensible input - and also by repetition, contextualized settings and experiential work. The fact that in many instances students can identify the oral and written input also leads to learning and, even more importantly, helps in understanding of the information and, thus, students feel motivated because they are able to communicate with one another and can limit the adjustment of their oral production while feeling that they are in a real communicative and safe in-group situation. In general, safety leads to taking risks in the L2 classroom so communicative acts become more challenging when students feel secure in their presentation to the rest of class. This secure situation also allows students to negotiate or provide alternatives, for example, those that can appear in a negotiation between a customer and client. Furthermore, the fact that this communication moves from pairs to the whole class means that the student has had the chance to repeat the same structures, vocabulary and speech acts a number of times before the general presentation. If this presentation is also supported by a presentation of how the data was obtained through a PowerPoint presentation the possibilities for quality feedback are high.

A second issue in language use is the appropriateness of the speech. As mentioned before, a large deal of the difficult or prefabricated patterns is immediately incorporated through the structure or vocabulary transfer between first and second language due to the students’ knowledge of their field. This transfer tends to be positive due to the fact that most input is constantly read on many websites (Levy et al., 1997; Levy et al., 1993; Garner & Degerman, 1967). For instance, any student who needs to work with lodging may have to go over the types of hotel facilities a number of times. This development is also reinforced continuously, especially by peer interaction (Gabbert et al., 1986) and because, in the end, the student tends to focus on certain information (i.e. whether the particular lodging has smoking rooms or not and so on). The teacher’s and students’ feedback or response to any student oral interaction facilitates student progress in the language (Bitchener, 2004) and gives both the protection and also the security to promote the move towards a higher stage of language development.
The repetition of the same patterns, structures, discourse markers and overall vocabulary based on students’ reading and the possibility of changing and playing with the same structures and lexical items refines student habits and routines and helps to combine them producing an unlimited number of different messages that will be necessary in real life. Therefore, the redundancy of certain structures will facilitate learning (Jamieson & Mewhort, 2005; Dienes et al., 1991). Working with similar websites is a genuine source of input, and cooperative learning assures repetition and feedback. Even the different rehearsals and peer interaction that the student goes through reinforce the contextualized use and serve to pay special and individualized attention to each piece of conversation (Lewandowsky et al., 2006; Hulme et al., 1991) and, consequently, leads from short-term comprehension to lasting long-term learning, from formational and phonological learning to output improvement (Baddeley et al., 1998; Adams & Gathercole, 1996; Service & Kohonen, 1995; Service & Craik, 1993; Service, 1992; Potts & Shiffrin, 1970). Thus the student will be able to incorporate what has been learned through the work with webquests into the speech. Additionally, the variety of input and topics provided by the use of webquests ensures a varied speech development. Even if accuracy in speech is thought to be a drawback because peer speech is usually less accurate than teacher talk, the student has nevertheless developed his or her communication skills and the feedback obtained through communication in real life is expected to have an influence in the changes that the student will make in the future and, eventually, speech will become more accurate.

### Student Production

The communication in English for Tourism in professional environments will tend to be specialized, functional and adequate to the professional context. Thus, it is the function of English for Tourism instructors to promote oral interactions to foresee and practice situations that students will probably face in travel agencies, hotels or tourism establishments. When students interact according to their findings in the webquest, they not only develop a repertoire of vocabulary and structures (as they would try to do in any traditional role play or classroom dialogue) but also need to give real information and solutions to a problem (necessary times, bookings, train class or cities to visit) in a communicative and realistic way with the possibility of unexpected variations (especially fares that can change from the beginning to the end of a booking).

Another relevant factor is the frequency of speech production. In fact, webquests without previous fixed direction allow for more frequent debate, comparison and student interaction. If a student needs to obtain the best or cheapest lodging or verify which of various possible companies quotes the best fares for a similar service, he or she will have to verify all the possible services but also discuss the best solution for the specific case with his or her peers. Therefore, the student will have to use the language more extensively but also without the same monitoring effect as would happen in a typical controlled class interaction.

### Webquests as a Source of Contextualization

The affective filter, according to various studies, can be lowered due to factors such as motivation, time and personal security (Nissani, 2003; Krashen 1982). Likewise, language acquisition is fostered if it occurs in a safe, supportive and, most importantly, contextualized way. Webquests are a good means of contextualizing the language because they have a direct link to the students’ field of specialization, they are fun, and provide students with new knowledge.

Groupwork is highly regarded when performed conveniently. Whereas English for Professional Purposes classes tend to be monotonous and vocabulary-oriented, cooperative environments are supportive and provide more opportunities for oral interaction than in many traditional classes. Besides, in this enriching atmosphere, the student takes language challenges that are unthinkable in larger classes where accuracy or drilling is the focus of instruction. According to the webquest structure presented in Figure 2 this atmosphere would be more secure for those students who may fear to speak because the pupils would only interact with the closest classmates at the beginning and the rest of the class when confidence is generated. Hence, in smaller groups, students would have more practice than in a whole class situation in which only very few usually have an active participation (usually those who are better skilled) but the weaker hardly get the chance to assume an active role. Although the implications of this change still need to be researched, an improved individual performance and a more participative attitude could be the consequence of more practice. This could be true even in high anxiety situations such as participation in whole-class
activities. For example, a student who is going to present a holiday package to the rest of the class would have had many previous rehearsal situations to promote security in the talk given in front of the other students of the class. As seen in Figure 3, feedback plays a significant role in this process. Students have the opportunity to rehearse their specific assignments and they have the general presentation rehearsal but, more importantly, they also have dialogues with peers so, by the time of the general presentation, they have done extensive individual and collective work with their classmates, received input, praise, and feedback and probably answered their (or the teacher’s) questions.

Another valuable asset not only to promote students’ professional skills and benefit motivation but also to improve their oral performance is the referential value of their speech. Students communicate real language in near-to-real situations and, indeed, with the intention of accomplishing real objectives. This also allows for spontaneity, negotiation and the need to solve problems that are unpredictable (for instance, if there are cancellations, lack of available places in transportation or fare changes between the beginning and the end of the booking). Therefore, students do not pretend but address a real-life situation.

Assessing Project-Based Work

As has been seen, the use of webquests in language learning has been addressed in many articles and professional publications. As in the example presented in the appendix, evaluation is an important element because project-based work requires seeing whether the objectives have been achieved and how. Assessment processes are very important in English for Tourism because prospective professionals are required to give presentations on final products (like the “Around the World Package”), so they can prove they possess the appropriate professional abilities, along with language skills to communicate, the ability to elaborate projects and perform in team-work. Furthermore, if self-esteem is to be developed in the prospective professional, it is necessary to make them aware of their own work and include personal and peer evaluations. Students can do that in the classroom and/or teacher’s office through reports,
teacher / learner-to-learner interviews, and in many other ways. In all cases, if oral development through the use of webquests is the objective, it is important that this evaluation should be oral or both written and oral. In this way, learners also need to report and use professional language and, additionally, metalanguage. They can assess their weaknesses and strengths along with elements that may need to be changed or readdressed in the future.

Conclusions

Although this paper addressed the use of webquests to develop oral competency in Tourism, the benefits and issues are common to many specialized fields and also to General English. The idea of integrating English for Tourism – CLIL – Project work needs to be emphasized in experiential learning. Although the researcher believes that webquests should not be seen as the main teaching and learning tool in English for Professional Purposes, it is clear that they can be a very valuable tool for providing students with many interaction opportunities in realistic settings, thus making the learning experience meaningful, experiential and very motivating. The approach hereby presented brings two types of benefits to the students: on the one hand, the learner’s professional competence in the use of both computers and Internet; on the other, the improvement of their language skills in aspects such as fluency (diminishing their anxiety in communicative situations), professional vocabulary increase and capacity to work using a foreign language. From a technical point of view, this approach permits the development of critical thinking in the choice of Internet sources promoting, at the same time, autonomous thinking by analyzing, contrasting and recommending prospective on-demand tailored actions (such as accommodation, transportation means, and so on) while receiving double feedback from the net itself and classmates. The language goals and the procedure are structured to facilitate the integration of productive (speaking and writing) as well as receptive (reading) skills but the program could also be supplemented by the inclusion of listening activities through the reproduction or incorporation of video repertoires such as those existing on many travel and tourism websites. These types of activities reinforce the application of the social constructivist theory more specifically since there is a clear relation between the students’ communicative and social interaction through critical thinking, dialogue and activity cooperation which reinforces individual, pair and group learning in the shape of an internet-based dialogue. It is in this way that students restructure their knowledge and communicative competence. This process must be observed by themselves and the instructor as a means of controlling, verifying and motivating language learning. Hence, evaluation has a significant value in the process.

Although the paper provides sound foundation in the topic, further research is necessary. In this sense, it will be interesting to see whether webquests have a real influence on longer-term learning and computer (and Internet) literacy, and how much conversation webquests may stimulate as measured by vocabulary and fluency. It would also be advisable to verify whether all the language skills required in this sort of project will be of use for the professional when he or she graduates and takes a real-life job. It would also be worth observing whether the language development obtained in the restricted spectrum of English for Tourism could be applicable to General English. The only problem that many higher institutions may have now is the availability of the necessary technological equipment but even this may no longer be a problem. In this respect, the extensive use of mobile devices such as mobile phones, PDA’s or i-PODs will surely be reflected in the increasing use of the Internet for communication and information in education and language learning. Further research on the topic may include: the study of teacher-student roles in this approach to webquests, the face-to-face interaction between students; the use of different technological devices for similar outcomes; the progress and willingness to speak more in a foreign language of students compared to that of students who do not follow the same approach but do analogue pair-driven exercises, professional inclusion or Internet for professional purposes development. Therefore, the field is still open to advances but especially so in terms of the application of new technology-driven approaches to foreign language teaching and learning, a field in which many foreign language teachers feel unfamiliar.

References


Appendix

Touring the World

Introduction

The global travel industry has become a major engine in the economies of many countries. Tourism has changed dramatically due to terrorism and fuel crises. Travel agents face the difficulty of finding combined travel at reasonable prices due to ignorance of foreign resources or companies and the difficulty of making bookings at an attractive cost. The García Travel Agency has picked your team to design and publish a combined holiday package around the world under the following conditions:

Continents to be visited: North America, Asia, Europe and Australia.

Transport to be used: train, plane, boat and bus.

Information about cities: those visited including lodging and food, and also tourist attractions, points of interest and places not to miss

Hotels: Do not use bookings through other travel agencies. 3-4 star.

Quote the final price and state full schedules, service features and comfort classes.

Task

Each group must gather information, and after discussion with other team members, choose the routes that will appeal to potential travellers (mostly independent). Each member of the group will research the means of transportation or local information or hotels in one continent.

Process

Find the appropriate web sites to do the bookings; find maps and places for your final oral report and presentation. Be sure and choose only the necessary means that will facilitate your work. Not every fact will be appropriate. Write a report that is clear, attractive and very persuasive. Because you will be producing a travel report, having one meeting and making a presentation, it is necessary for you to keep notes of your information or maintain records of your research.

Resources

Find your own resources (the teacher will be available if necessary). However, keep in mind (only for emergency reasons) that Best Western is the largest hotel chain in the world, that Amtrack is the national railways company in the US, and so on.

Evaluation

The following rubric will be used to evaluate you. Students will be evaluated on individual research components, their presentation and meeting, their final report and a justified self-evaluation.

<table>
<thead>
<tr>
<th>Does Not Meet (0-4 points)</th>
<th>Meets (4.5-7)</th>
<th>Excels (7.5-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Participation</td>
<td>All members participate equally</td>
<td>Not have equal participation</td>
</tr>
<tr>
<td>Training skills</td>
<td>Group shows lack of skills in the tasks and does not give a good presentation</td>
<td>Some group members lack skills in the tasks or perform below average in the general presentation</td>
</tr>
<tr>
<td>Creativity</td>
<td>Group shows no creativity and seems unprepared and unenthusiastic</td>
<td>Group shows some creativity but little enthusiasm and preparation</td>
</tr>
<tr>
<td>Finished Product</td>
<td>Project not complete</td>
<td>Project is completed but lacks understanding of the tasks undertaken in the project / webquest</td>
</tr>
</tbody>
</table>

*Figure 4: Grading system*

**Conclusion**

You have done extensive research and learned much about the international travel industry. You will have experienced the difficulty of working on your own but also having to work cooperatively. In your cooperative groups you were required to make difficult decisions about which possibilities were optimal. Your findings, decisions and analysis produced a report which had to be contrasted in a meeting and presented. Do you want to be the first to work? Well, if you ever become a travel agent you will be prepared and know what to do.

**Web support**

Although for practical reasons this exercise did not provide the web addresses, a guided webquest could include at least some of the following recommendations if teachers want to provide their students with some help.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Websites and their use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization of meetings and conventions</td>
<td><a href="http://corporate.iexplore.com/corporate/index.jhtml">http://corporate.iexplore.com/corporate/index.jhtml</a> (Business travel)  &lt;br&gt;<a href="http://www.gulfshores.com/conventions-meetings/conventions/services.asp">http://www.gulfshores.com/conventions-meetings/conventions/services.asp</a> (Meetings and conventions)</td>
</tr>
</tbody>
</table>
The effects of web-based instruction navigation modes on undergraduates’ learning outcomes

Yu-Chang Hsu1*, Huifen Lin2, Yu-Hui Ching1 and Francis M. Dwyer1

1Instructional Systems Program, College of Education, Pennsylvania State University, 315 Keller Building, University Park, PA 16802, USA // Tel: (814) 865-0473 // Fax: (814) 865-0128 //
2English Department, National Kaohsiung Normal University, Taiwan, 1F, 33, Lane 400, Sec.3, Mincyuan Rd., Tainan, Taiwan // Tel: 886-6-3500881 // Fax: 886-6-2050995 //
yzh109@psu.edu // huifen5612@yahoo.com.tw // yuc148@psu.edu // fmd@psu.edu

*Corresponding author

ABSTRACT

The purpose of this study was to examine whether matching navigation mode of a learning environment with learners’ preferred navigation mode would facilitate their learning in a web environment. Sixty-eight undergraduate students were randomly assigned to treatments (linear vs. nonlinear navigation mode) and received four criterion tests designed to measure different educational objectives immediately after interacting with the instructional material about human heart. The results suggested that matching or mismatching navigation mode of the learning environment with learners’ preference did not lead to significant differences in learning outcomes. However, there were significant differences in achievement between groups with different navigation mode preferences. Our findings indicated that students preferring nonlinear navigation had significantly higher achievement scores on higher levels of learning outcomes. Based on the results of the study, we discussed specific recommendations for instructional designers and the issue concerning matching/mismatching navigation modes with learning preferences.

Keywords

Navigation mode, Undergraduate students, Web-based instruction, Matching/mismatching, Navigation preference

Introduction

Web-based instruction (WBI) has gained considerable popularity in education due to its benefits such as allowing learner control (Laurillard, 1993; Chen, 2002; Alomyan, 2004), providing practice of self-discipline, time-management (Daugherty & Funke, 1998), and 24/7 accessibility (Chuang, 2000). To maximize the educational value of WBI, researchers in this field have been exploring the design and usability of websites. Typical design considerations of instructional websites include navigation tools, response time, credibility, and content (Nielsen, 2000). Among the above considerations, the navigation mode applied in designing navigation tools is an important element since it determines how learners experience the information they need to acquire. It is asserted that the nonlinear capability of the World Wide Web (WWW) has great potential for education because of the opportunities it offers individuals to control their own learning (e.g., Chen, 2002; Alomyan, 2004). Being able to determine what and how to learn also makes the learning experience meaningful to individual learners (Laurillard, 1993). However, some researchers (e.g., Lazonder, Biemans, & Wopereis, 2000) argued that certain learners are not prepared to construct their own learning paths. The question of interest is, “Do all the learners today benefit from the web’s nonlinear capability? If not, which types of individuals would or would not benefit from this capability?” Through examining the interaction between web navigation modes and learners’ characteristics regarding web navigation, useful suggestions for designing web-based instruction were provided based on the results of this study.

Literature review

Nonlinearity vs. Linearity

Web-based instruction (WBI), taking advantage of the advancement of hypermedia technology, has become a popular alternative to mainstream face-to-face instruction. The potential of WBI has led to the shift of research focus to examining variables that may contribute to the success of this kind of instruction. Research has shown that the
advantage of WBI lies in its affordance of nonlinear interaction (Laurillard, 1993). Nonlinearity of WBI is believed to provide individuals with learning decisions that allow them to control their sequence and pace while learning the target material. In addition, allowing individuals to have control over their learning also makes them more motivated to learn (Keller, 1983).

In a hypermedia/hypertext learning environment, information is organized by various nodes and links. Texts, visuals, or organizational cues are displayed by nodes, while links are used to connect the nodes (Barab, Young, & Wang, 1999). The initial appearance of these nodes and links constitutes a mental cognitive map in learners in terms of the scope and structure of the information to be learned. Linear navigation ensures that the decision relative to information processing is in the hands of the instructional designer rather than the learners. Research has shown that by presenting information more explicitly and structurally, linear navigation presentation makes the material more likely to be assimilated by learners (Dillon & Gabbard, 1998; Meehan & Shubin, 1997).

**Hypermedia navigation preference**

A large body of research has investigated how individual differences affect students’ learning in hypermedia/hypertext environments (e.g., web-based instructional modules) that adopted varied navigation patterns. The individual differences explored include learners’ cognitive styles (e.g., Dufresne & Turcotte, 1997; Chen, 2002; Gauss & Urbas, 2003), prior knowledge (e.g., Recker, Ram, Shikona, Li, & Stasko, 1995; McDonald & Stevenson, 1998; Lawless & Kulikovich, 1998; Holscher & Strube, 2000; Last, O’Donnell, & Kelly, 2001), gender difference (e.g., Ford & Miller, 1996; Reed & Oughton, 1997), learner interest/motivation (e.g., Lawless & Brown, 1997; Chou & Liu, 2005), and learners’ computer expertise/experience (e.g., Calcaterra, Antonietti, & Underwood, 2005).

In their review examining research on developing a hypermedia learning model, Chen and Macredie (2002) found that cognitive styles played a critical role in affecting students’ learning in hypermedia systems. As they also pointed out, *Field Dependence* proposed by Witkin, Moore, Goodenough, and Cox (1977) was the most widely studied example of dimensions of cognitive style—it helps explain learners’ preferences in terms of organizing and representing information. According to Jonassen and Grabowski (1993), when encountering information, field dependent (FD) persons tend to accept the existing structure of incoming information, while field independent (FI) persons are more likely to reorganize or restructure the information. Since FIs are more likely and able to impose their own structure, they are more tolerant with ambiguity and uncertainty in learning environments (e.g., nonlinearity in web-based instruction) than FDs. In their study of navigation in hypermedia learning environments, Dufresne and Turcotte (1997) indicated that both FD and FI students found it easier to navigate in linear than in nonlinear environments, given that a linear format was more appropriate for presenting complicated content. However, FD students felt more disoriented than their FI peers, and had problems understanding and remembering information they encountered in the nonlinear environment.

**Matching versus mismatching**

The era of hypermedia and the development of advanced technology have empowered web-based instructional designers to design learning materials/environments that would optimize students’ learning by accommodating various learner differences. Instructional websites with such capabilities were encouraging and attractive in their own right, but the extent to which the custom design can actually be tailored to meet learners’ needs remains unclear. Most importantly, learning effectiveness that can be attributed to the design is still open to questions. Past research has suggested that presenting knowledge to learners with an expert’s knowledge structure in a hypertext system seems to be beneficial in assisting learners in processing and acquiring that knowledge (Shapiro & Niederhauser, 2003; Su & Klein, 2006). However, an inflexible mapping of expert knowledge structure on learners did not always lead to desirable learning outcome in learning contexts that did not account for individual differences, background, prior knowledge, and other characteristics that usually accompany learning effectiveness and efficiency. On the other hand, learning style (used interchangeable with cognitive style in this paper) and experience also entail problems when the learning environment allows students to proceed at their own pace. Learners might not be experienced enough to choose the most effective routes to go about learning the targeted knowledge (Gall & Hannafin, 1994; Su & Klein, 2006).
Until today, it is still a dilemma for instructional designers creating web-based instruction regarding whether or not the presentation of instructional material should be in accordance with students’ learning styles. Although researchers and educational practitioners have urged and promoted the ideas of incorporating differences/preferences of individual users into instructional design tasks, the core issue concerning matching or mismatching of learning and instructional presentation styles should be given top priority over other instructional decision-making considerations.

Pask’s work dated back to the 1970s remained to be the most consistent and systematic exploration of the topic concerning matching or mismatching of instructional presentation style and students’ identified cognitive styles. In a detailed review of Pask and his associates’ studies, Ford (2000) indicated that Pask et al. studies were important because they found consistent and dramatic effects of learning performance that strongly favored the matching of presentation style of material with learners’ cognitive style. Pask (1976) identified two approaches that people generally adopt in the learning processes, i.e., holists and serialists. The former learners are more inclined to take a global approach to learning. They 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. Serialists, on the other hand, would tackle individual details first, or take so-called “local learning approach” (Ford, 2000, p. 543), connect the separate topics, and finally form the overall picture.

Using complex academic subject matter, including biological taxonomies, the operon, reaction kinetics, and Henry VIII’s reign, and two complex as well as learner-dooming tests (i.e., free-learning and teach-back to identify learning styles), Pask’s studies found high correlation between matching and mismatching of teaching style and learning style. Pask (1976) concluded that people tended to “… consistently prefer a particular type of learning strategy…” if given a choice and that “… if the teaching strategy is matched to the same type of learning style… the student will learn more quickly and retain the information for longer” (p.132). On the other hand, a mismatched condition tends to result in unsatisfactory performance.

Ford and Chen (2001) explored the relationship between matching and mismatching of instructional presentation styles with students’ cognitive styles. The presentation styles included breadth first (a Holist strategy) and depth first (a Serialist strategy); cognitive styles included field dependence and field independence. In their matched conditions, field-independent students worked with the depth-first version of the learning materials while field-dependent students worked with the breadth-first version. The participants in their study were 73 postgraduate students engaging in a task of creating Web pages using HTML. The results suggested that the matched-conditions group had better performance than the mismatched-conditions group only for male students. To some extent, this study provided support for the effect of matching condition on learning outcomes.

Ford (1985) conducted a study to explore if experienced and successful learners, e.g., postgraduate students, identified as having holist or serialist learning style can learn equally well from material designed specifically to suit serialist and holists respectively. Contrary to Pask’s instruments used to identify learning style, Ford used a quick and easy questionnaire measure termed “Study Preference Questionnaire.” The results agreed with Pask’s work. Students, even though as experienced as the ones in Ford’s study, cannot be expected to learn equally well from material designed to suit holists or serialists respectively if their learning styles were contrary to the instructional presentation style.

In another study, Ford (1995) investigated the relationship between levels and types of mediation in instructional material and individual differences. Thirty-eight participants were first tested for field dependence/independence using Cognitive Style Analysis (CSA) and then learned a computerized version of Pask and Scott’s instructional materials designed to suit either holist or serialist learning strategies. The study found that students in matched conditions performed significantly better than those in mismatched conditions, consistent for both holists and serialists.

In this study, we focused on examining whether matching navigation mode of the learning environment with learners’ preferred navigation mode would facilitate their learning in a web environment. The preferred navigation mode was used as an indicator of learners’ cognitive style in this research. We selected undergraduate students as participants due to the increasing number of web-based courses for undergraduate students, hence the need to provide instructional design suggestions for web-based courses for this specific population. Additionally, the instructional material used in this study was enriched with animation since it is becoming popular for its capability of enhancing learning in today’s web-based instruction.
Specifically, this study aimed to answer the following research questions:

1. Is there a difference between matching and mismatching assigned navigation modes in web-based instruction with learners’ preferred modes on their learning outcomes?
2. What is the effect of assigned navigation mode (linear vs. nonlinear) of web-based instruction on learning outcomes?
3. What is the effect of learners’ preferred navigation mode (linear vs. nonlinear) on their learning outcomes?

**Methods**

**Research Design**

This research employed a two-way MANCOVA design. The two independent variables were assigned navigation mode (nonlinear vs. linear) and preferred navigation mode (nonlinear vs. linear). Four criterion measures – drawing test, identification test, terminology test, and comprehension test – were utilized to assess the effect of the intervention on the different criterion measures. Participants’ prior knowledge of physiology, measured with a 36-item test, was used as a covariate to control for its effect on learning performance. Figure 1 presents a graphic illustration of the experimental design.

![Figure 1. Research design](image)

**Instructional material**

The instructional material used in the study was an 1821-word unit on the parts of the heart and its functions (Dwyer, 1978). The original materials were developed into a web-based instructional format accompanied by animated visuals to facilitate students’ learning of the material. Various animation techniques were used to provide dynamic illustration of concepts that were difficult to comprehend. Some examples are progressive reveal, motion, pop-in verbal, motion, contraction, and expansion. Students were allowed to review the animated visuals as many times as needed by clicking on the “Play Animation” button.
The instructional module consisted of five units: 1) The heart’s structure; 2) The veins and arteries; 3) The valves of the heart; 4) The blood flow through the heart and 5) The phases of the heart cycle. Content for each unit was presented in several subunits, with one subunit per webpage. There were a total of 20 subunits for the entire instructional module. Animations were displayed on the right-hand side of the screen to complement the text to their left. The respective navigation treatments employed in the study are described in the following section.

Assigned navigation mode

Linear navigation

A navigation menu with hyperlinks showing “back” and “next” was placed beneath the major content text on each webpage. This menu was the only tool that learners could utilize to proceed in the web-based instruction. The navigation design was extremely linear in that learners could only proceed in a predetermined way. The major unit titles on the left side of the screen indicated students’ current progress in this module. Moreover, only unit titles were presented along with corresponding content text. The subunit titles for each unit were not presented for more detailed overview as were in the nonlinear mode (see below). This difference was meant to be a difference between both modes, since the design of the linear module was intended to keep learners focusing on immediately available content on each page. Figure 2 provides a screenshot of the linear-navigation treatment.

Nonlinear navigation

The instructional content in the nonlinear navigation treatment is exactly the same as that in the linear version. However, the navigation menu beneath the text, which allows students to move back and forward, was eliminated (see Figure 3). Instead, this nonlinear navigation design allows learners to navigate the 20-page instructional material at their own pace and in unarranged sequence using an interactive dropdown menu on the left side of each webpage. The treatment was designed to be in great contrast with the liner counterpart. In this nonlinear version, the participants have immediate access to any webpage at any time by placing the cursor over any one of the five major units that display automatically the titles of the subunits (see Figure 4). To access a sub-unit of their choice, learners could mouse over and click the sub-unit title, which then brought them to the desired unit.
Navigation preference measures

Instead of identifying students’ cognitive style on the field dependence dimension that is related to individuals’ preferences in navigation, a 10-item navigation preference measure was adopted to directly identify participants’ preferences for specific web navigation modes (Hsu, 2006). In this measure, there were 10 pairs of websites (i.e., one pair per item) for students’ review. Each pair of websites included both linear and nonlinear navigation material that contained the identical content and subcategories, with the only difference being navigation mode (i.e., linear vs.
nonlinear). Each module was labeled as either A or B respectively on top right corner of the page (e.g., Figure 5 and Figure 6).

**Figure 5.** Navigation preference: Screenshot of a linear navigational webpage

**Figure 6.** Navigation preference: Screenshot of a nonlinear navigational webpage

On each “choice” page (e.g., Figure 7), the participants were asked to click on a radio button to indicate their navigation preference immediately after they viewed each pair of navigation prototypes. It has to be noted that the letter A or B is not associated with a specific navigation mode and was used randomly to label either linear or nonlinear presentation.
A scoring scheme was employed to classify participants into two different navigation preferences. Using the preference for nonlinear navigation as the coding standard, participants who identified their preferences for nonlinear navigation in 6 or more pairs (i.e., possible score: 6 to 10) were categorized as “Preferring Nonlinear Navigation,” while those who identified their preference for nonlinear navigation in 5 or fewer pairs (i.e., possible score: 0 to 5) were categorized as “Preferring Linear Navigation.” This navigation preference measure has a high reliability (KR-20 = .94), that is, participants who preferred a specific navigation mode tended to consistently select that mode. In fact, only 3 of the 68 participants scored 5 (i.e., middle range), others scored either 4 and below or 7 and above. Also, it is worth mentioning that the KR-20 reliability values reported in this paper were calculated based on the participants’ scores on each measure/test in this study.

**Instruments**

The study employed four tests, drawing, identification, terminology and comprehension test, to assess treatment effects, and a physiology test to evaluate students’ prior knowledge related to the learning content. The instruments were evaluated by a panel of content experts when they were developed in the 1960s and had undergone several revisions. The current version of instruments has been used in more than hundreds of experimental studies (Canelos, 1987), systematically manipulated making use of other independent variables related to instruction, and facilitated systematic explorations of visualized web-based learning environments. The following is a description of each individual test.

**Physiology prior knowledge test**

A 36-item test (KR-20 = .63), validated by content experts, was presented in a multiple-choice format to evaluate students’ prior knowledge of general physiology (Dwyer, 1978). The purpose of using this test was to control for the effect of the prior knowledge related to the content of the instructional material on students’ posttest performance.

**Criterion measures**

Four criterion measures, also developed by Dwyer (1978) and validated by content experts, were used to assess learning outcomes of the instructional treatment designed specifically for this research. The four tests were: 1)
drawing test (KR-20 = .93); 2) identification test (KR-20 = .87); 3) terminology test (KR-20 = .87); 4) and comprehension test (KR-20 = .84). The total score for each test is 20. A composite score (KR-20 = .96) adding scores of the aforementioned four scores was used to measure each student’s overall achievement on learning from the instructional material. The four tests are briefly described as follows.

The drawing test

This test was designed to evaluate student ability to construct and/or reproduce items in their appropriate context. This test provided the students with a numbered list of terms corresponding to the parts of the heart discussed in the instructional presentation. The students were required to draw a representative diagram of the heart and place the numbers of the listed parts in their respective positions. For this test the emphasis was on the correct positioning of the verbal symbols with respect to one another and with respect to their concrete referents.

The identification test

The test was designed to evaluate students’ ability to identify parts or positions of an object. This multiple-choice test required students to identify the numbered parts on a detailed drawing of a heart. Each part of the heart, which had been discussed in the presentation, was numbered on a drawing. The objective of this test was to measure students’ ability in using visual cues to discriminate one structure of the heart from another and to associate specific parts of the heart with their proper names.

The terminology test

This test consisted of items designed to measure knowledge of specific facts, terms, and definitions. The objectives measured by this type of test are appropriate for all content areas, while assuming an understanding of the basic elements is a prerequisite to the learning of concepts, rules, and principles.

The comprehension test

Given the location of certain parts of the heart at a particular moment of its functioning, the students were asked to determine the position of other specified parts of the heart that are functioning at the same time. This test required that the students have a thorough understanding of the heart, its parts, its internal functioning, and the simultaneous processes occurring during the systolic and diastolic phases. The comprehension test was designed to measure a type of understanding in which the individual can use the information being received to explain some other phenomenon.

Participants

Convenient sampling was used in this study. Sixty-eight (N = 68) undergraduate students were recruited from an entry-level statistics class from a northeastern research university in the U.S. to participate in this study. Among the 68 students, there were 25 males and 43 females, while 86.7% of them were 19 to 20 years old. The students received extra class credit for their participation.

Results

The average prior knowledge test score of the 68 students was 21.12 (SD = 4.08) out of the total possible score of 36, with the lowest score of 11 and the highest score of 32. As for the students’ preference for navigation modes, 56 students were identified as preferring nonlinear navigation (preference score: M = 9.61, SD = 0.89), while 12 as preferring linear navigation (preference score: M = 2.33, SD = 2.31). Table 1 shows the descriptive statistics of assigned and preferred navigation modes on the criterion measures.
Table 1. The learning outcome means by navigation modes

<table>
<thead>
<tr>
<th>Criterion Measure</th>
<th>Assigned Navigation</th>
<th>Preferred Navigation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonlinear (n = 27)</td>
<td>Linear (n = 7)</td>
</tr>
<tr>
<td>Drawing</td>
<td>13.11 (5.47)</td>
<td>11.29 (7.70)</td>
</tr>
<tr>
<td>Identification</td>
<td>16.11 (3.82)</td>
<td>14.29 (5.19)</td>
</tr>
<tr>
<td>Terminology</td>
<td>13.67 (4.54)</td>
<td>10.57 (5.59)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>11.59 (3.96)</td>
<td>8.14 (3.63)</td>
</tr>
<tr>
<td>Total Score</td>
<td>54.48 (15.86)</td>
<td>44.29 (20.63)</td>
</tr>
</tbody>
</table>

|                   | Nonlinear (n = 29)  | Linear (n = 5)       |
| Drawing           | 13.31 (4.90)        | 8.00 (9.38)          |
| Identification    | 15.34 (4.64)        | 12.60 (5.08)         |
| Terminology       | 13.21 (5.24)        | 10.20 (5.36)         |
| Comprehension     | 12.07 (4.47)        | 6.00 (6.36)          |
| Total Score       | 53.93 (17.62)       | 36.80 (24.67)        |

* the maximum score for each criterion measure is 20.

We further analyzed learning outcomes by comparing participants who received treatments either suited or contrary to their preferred web navigation style. The analysis revealed that 32 participants received navigation tools suited to their preferred navigation style (i.e., matched group: nonlinear-preference students receiving nonlinear treatment; linear-preference students receiving linear treatment), and 36 students were assigned navigation tools contrary to their learning preference (i.e., mismatched group: nonlinear preference students receiving linear treatment; linear preference students receiving nonlinear treatment). Table 2 shows the means and standard deviations for each criterion measure of the matched and mismatched group.

Table 2. The learning outcome means by matching groups

<table>
<thead>
<tr>
<th>Criterion Measure</th>
<th>Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Matched (n = 32)</td>
</tr>
<tr>
<td>Drawing</td>
<td>12.31 (6.33)</td>
</tr>
<tr>
<td>Identification</td>
<td>15.56 (4.15)</td>
</tr>
<tr>
<td>Terminology</td>
<td>13.13 (4.76)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>10.72 (4.76)</td>
</tr>
<tr>
<td>Total Score</td>
<td>51.72 (18.22)</td>
</tr>
</tbody>
</table>

The omnibus test of one-way MANCOVA, using “matching” as the independent variable and controlling for prior knowledge of general physiology, indicates there was no significant difference between the matched and mismatched groups on the learning outcomes ($F(4,62) = 1.15, p = .344$).

We further conducted a two-way MANCOVA, controlling for prior knowledge of general physiology, to investigate the interaction between assigned navigation mode and preferred navigation mode and found that it was not statistically significant at the .05 level ($F(4,60) = 1.255, p = .298$). As a result, the main effects of respective assigned navigation modes and preferred navigation modes were examined. The results indicated that students’ learning outcomes was not significant for the assigned navigation mode at the .05 level ($F(4, 60) = 2.94, p = .028$), but significant at .05 level for the preferred navigation mode ($F(4,60) = 2.94, p = .028$).

Table 3. The learning outcome means by preferred navigation modes

<table>
<thead>
<tr>
<th>Preferred Navigation (N)</th>
<th>Drawing</th>
<th>Identification</th>
<th>Terminology</th>
<th>Comprehension</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonlinear (n = 56)</td>
<td>13.21</td>
<td>15.71</td>
<td>13.43</td>
<td>11.84</td>
<td>54.20</td>
</tr>
<tr>
<td>(5.14)</td>
<td>(4.24)</td>
<td>(4.88)</td>
<td>(4.20)</td>
<td>(16.64)</td>
<td></td>
</tr>
<tr>
<td>Linear (n = 12)</td>
<td>9.92 (8.20)</td>
<td>13.58 (4.98)</td>
<td>10.42 (5.25)</td>
<td>7.25 (4.81)</td>
<td>41.17</td>
</tr>
</tbody>
</table>

* the maximum score for each criterion measure is 20.

Since the MANCOVA test revealed the main effect of preferred navigation modes on the students’ learning outcomes, the dependent variables were then considered separately. The only difference reaching statistical
significance was the comprehension test \((F(1,63) = 10.856, p = .002)\), with the nonlinear preference group receiving higher average scores \((M = 11.84, SD = 4.20)\) than the linear preference group \((M = 7.25, SD = 4.81)\). Table 3 shows the means and standard deviations of scores obtained by students identified as preferring different navigation modes.

**Discussion and conclusion**

Our study revisited the effects of matching/mismatching of cognitive styles and presentation styles on learning by investigating the matching/mismatching of one’s preferred navigation mode (arguably indicating one’s cognitive style) and assigned navigation mode in the web-based learning environment. Students were assigned to one of the two treatment groups—a linear navigation group or a nonlinear navigation group, after taking a test on their website navigational preferences. Treatment effects were assessed with four criterion tests that address lower level factual/descriptive knowledge as well as more complex intellectual skills.

Results indicated there was no significant difference in learning performance between the matched and mismatched groups in all four criterion measures. That is, we could not conclude that matching navigation mode of web-based instruction with one’s preferred navigation mode provide participants advantages in learning over their counterparts receiving mismatched navigation mode. The findings of our research did not support those in past research studying the effect of matching (i.e., Pask & Scott, 1972; Pask, 1976, Ford 1985; Ford, 1995) which suggested learning performance in matched conditions was significantly superior to that in mismatched conditions. One plausible reason for our finding might be that “… human beings are highly adaptable it may be possible for an individual with any sort of competence to learn, in the end, according to any teaching strategy…” as indicated by Pask and Scott (1972, p. 221), although Pask et al. also emphasized that “… the rate, quality and durability of learning…” might depend on whether or not the teaching strategy matches the individual. Our finding suggested that matching instructional style to learners’ preferred learning style specifically in navigation tools in a web-based environment might not be a critical issue that needs to be taken into design consideration, if the aim of instruction is to facilitate learning achievement.

Moreover, our criterion tests, assessing knowledge ranging from facts to comprehension, were in a format of multiple-choice questions. Assessment of this kind was quick and easy to administer and score; nonetheless, they might not be sufficiently discriminative to reveal differences in learning effectiveness from matched and mismatched conditions. One probable reason for the difference in results between ours and those in Pask was that students in Pask’s study were required to go through the program repeatedly until an error-free run had been achieved before they can take the tests, i.e., students accessed the program in a way similar to programmed instruction. In our study, we did not impose such mechanism, and students were, instead, encouraged but not required to review the material to a certain degree of mastery before they proceeded to take the tests.

Furthermore, our results indicated that linear and nonlinear navigation modes had equal effects on learning outcomes, which concurred with Dufresne and Turcotte’s (1997) findings, but contradicted those of Baylor’s (2001). Baylor found that nonlinear navigation was more effective in facilitating task performance in the form of example and main point generation, and that linear navigation actually led to more disorientation than nonlinear navigation. The result of our study suggested that the expected differential effects of nonlinearity and linearity navigation tools were not distinctive and therefore might not be an important consideration in web-based instructional design. It is far from clear why the nonlinear approach, as preferred by most students in our study, did not result in more learning gain when compared to the linear version. However, it must be noted that the authors did not collect the en-route of students’ learning process. The experimental material used in the study was developed into an online version from an original text-based booklet that was designed to be studied in a linear mode of progression. It was possible that, students receiving the nonlinear navigation might have accessed the material in a predominantly linear fashion as the linear students did. There might not be expected differences in achievement should both groups of students study the material in the same or similar way.

The most astounding finding of our study was that learners preferring nonlinear navigation had significantly better performance on relatively higher order learning objective (i.e., comprehension measure in this research) than their counterparts preferring linear navigation. This finding was contrary to Wang and Beasley’s (2002) study in which they found hypermedia preferences did not have a significant main effect on cyber-students’ task performance in web-based learning environment. While there is no direct evidence suggesting that nonlinear preference learners are
better learners of intellectually more challenging task from our study, it seems plausible to argue that these students might share certain characteristics of FI individuals, considering FIs’ preference for nonlinear navigation. As Jonassen and Grabowski (1993) pointed out, FIs are more likely than FDs to “… reorganize, restructure, or represent information to suit their needs, conceptions, or perceptions” (p. 87), it would also be fair to argue that FIs might engage in deeper information processing and deeper learning, hence having better learning performance. Previous studies did suggest that FIs performed better than FDs on tests measuring different educational objectives (Dwyer & Moore, 1995) in certain subject domains, for example, mathematics, natural science, and social science (Tinajero & Paramo, 1997), and on technical courses in an information management program (Murphy, Casey, Day, & Young, 1997). With the characteristics shared by FIs and those preferring nonlinear navigation, it is likely that nonlinear preference individuals engage in deeper information processing during learning, hence obtaining better performance, especially on relatively higher order cognitive objective that requires deeper information processing.

Educational implications and suggestions for future research

The implications of our study could be reported and interpreted as follows:

1. Presenting material indiscriminately to a class of mixed learning preferences regarding navigational tools provided is warranted at least to the extent that those received material that is not matched with their navigation preference would not perform any worse than those receiving matched material.

2. Navigation tools, either linear or nonlinear, led to equal learning effects in a web-based learning environment, suggesting that navigation mode might not be a critical factor for consideration while designing web-based instruction. Other factors such as cost of material development and learning efficiency concerning time-on-task might be priorities instead.

3. Students favoring nonlinear navigation have a tendency to achieve higher-order learning objectives regardless of the presentation format of the material. Special attention must then be paid to those preferring linear navigation. Enhancement strategies might need to be embedded in the linear material to scaffold and in turn, improve learning.

This study has reactivated the discussion concerning whether the matching or mismatching of instructional presentation to student cognitive style would result in improved performance. Additionally, this research also expands our understanding of such issue, in particular, in an enriched online environment with animation. As the research findings were inconsistent with those in some of the past studies, more conclusive and consistent evidence is needed to determine the roles of matching instructional presentation with students’ cognitive style, and the extent to which they impact learning (e.g., students’ motivation, time on task), in order to design the online environment more effectively.

However, it should be noted that qualitative data on students’ navigation patterns were not collected in this study. Muller, Hobbs, and Moore (2002) have indicated that examining a learner’s navigation patterns could inform us their ability to perform specific kinds of tasks. More research complemented with qualitative analysis is encouraged to help develop a deeper understanding of the interactions between learners and teaching material, in which other factors such as gender, prior knowledge, forms of learning, and subject domains may interact to various degree to affect learning (Ford & Chen, 2001).

Meanwhile, our study included only 68 participants and the sizes of navigation preference groups were unequal, which might have limited the generalizability of the research findings. It is suggested that larger scale studies be conducted to provide additional empirical evidence. Moreover, merely two simplified versions of navigation modes were included in the web-based instructional modules in this study. They should be viewed as examples of contrasted navigation modes rather than the norm of websites. Future research could be conducted to examine various types of website organizations/structures regarding the effects of assigned and preferred web navigation modes on different levels of learning outcomes.

References


Information society needs of managers in a large governmental organisation

Elizabeth Broos
Faculty of Education: University of Pretoria, South Africa // ellybroos@yahoo.co.uk

Johannes C Cronjé
(Corresponding author) Faculty of Informatics and Design, Cape Peninsula University of Technology, South Africa // cronjej@cput.ac.za // Tel +27825585311 // Fax +27214603553

ABSTRACT
Dealing effectively with information and communication technology in the information society is a complex task and the human dimension is often under-estimated. This paper tries to give a voice to some managers about their experiences with information, communication and technology in their working environment, which involves participating in a learning organisation, knowledge management and communities of practice, competency management, ICT-Security awareness management as well as and innovation and change management. Managers of a large governmental organisation in the Netherlands were polled in a questionnaire requesting their responses in using ICT as well as involvement in the above categories. The responses of 246 of them were analysed. It was found that they knew full well that their strongest need was for a conceptual understanding of the implications of ICT in their changed work environment. This need overarched the elementary end-user training needs of what buttons to press to achieve what end. The research findings emphasize that as education is increasingly need-driven instead of content-driven it is necessary to reconsider the curricula of higher educational institutes especially with regards to management training. Furthermore, this research highlights a need for on the job training and performance support for middle managers.

Keywords
ICT skills, Digital work environment, Information society, Learning organisation, E-leader

Introduction
The organisation, the technology, the employees as well as the information society each play an important role in dealing effectively with information and communication technology and they influence each other. Especially managers play an important role in firstly implementing new technologies in the working environment and secondly in motivating their staff to use the information, communication and the technology effectively. Yet managers often struggle to define their role in the new work situation. In this paper some answers are given to the question: What are the reported experiences of managers regarding information and communication technology in their working environment? Working from the perspective of a person who is responsible for facilitating learning in a large governmental organisation in Europe, we wanted to know two things. Firstly, what ICT skills do they already have that we may draw upon when creating a culture of learning in an organisation? Secondly what are the ICT skills required by middle managers?

Literature shows that managers play an essential role in managing information as well as implementing information and communication technology in organisations (Hargrove, 2001; Kluytmans, 2005; Boonstra, 2005). This is a complex process for which managers are often not sufficiently prepared (Beijen, Broos & Lucas, 2003). Hence it appears important to research the area of information and communication technology in the work environment further from the perspective of managers in order to identify needs for further development as well as for performance support in this regard. Yukl (2006) also calls for research in which more insight is obtained in how managers deal with information and communication technology in the working environment. Furthermore, Davenport & Prusak (1997) argue that in order to optimize any information system, it is important to investigate the human interactions with the system as well as their interpretation of the system in the context of their working environment. This leads us to the important question if managers are indeed ready and prepared to deal effectively with the adjusted ways of working required in the information society by using information and communication technology. Such information will be valuable when determining the ICT literacy requirements for management training, as well as determining the extent to which ICT would be a suitable conduit for management education.
Literature survey

Feather (2004, P. 209) states that “the information society can perhaps best be understood as a society that has developed information technology and is learning to use it.”

The technology in the information society has made the current level of globalization and the sharing of information possible through international networks and mobile technologies (Boonstra, 2005; Hargrove, 2001; Feather, 2004) and the amount of available information is increasing exponentially. Managers are confronted with the task of implementing changes and innovations as organisations continue to adjust to the requirements of the information society (Hargrove, 2001; Boonstra, 2005). Technology makes alternative ways of working possible and the requirements of the information society are such that implementing alternative ways of working are necessary to adjust to the changed requirements. Hargrove (2001) has noticed that managers could play a crucial role in this regard on three different levels: firstly in influencing the strategy of the organisation in this regard, secondly in implementing the new strategy and thirdly in influencing their subordinates in participating in the new ways of working and using ICT. However there are also authors that warn that the technology of the information society can never replace the social networks and resources that make learning and working possible (Brown & Duguid, 2000).

Our research identified the following components of adjusted ways of working and leading for managers in the information society (Broos, 2007):

- Participating in a learning organisation, knowledge management and communities of practice
- Competency management
- ICT-Security awareness management
- Innovation and change management

In the following section the literature pertaining to each of the components is discussed.

Participating in the learning organisation, knowledge management and communities of practice

Organisations need to become adaptive and flexible (Belasen, 2000) and it is important that they change into learning organisations, in order to keep up with continuing changes required in the information society (Wenger, 2000; Senge, 1990). A corporate curriculum has become very important in such organisations, according to Kessels & Keursten (2001) and working and learning come together in the work environment. Just-in-time learning and mobile technology could also play a role in the learning organisation (Hargrove, 2001; Traxler, 2005).

The importance of a positive attitude towards learning and life-long learning in the information society is generally accepted, but Hargrove (2001) takes this even further and claims that new ways of learning are required in the information society. He calls this learning process transformational learning in which learners fundamentally change in the way they think and behave in their working context. Other authors support this notion (Yukl, 2006; Zaccaro et al., 2006). Individuals can thus learn in an organisation, but the learning capacity of an organisation can also be increased by implementing a knowledge infrastructure and dealing with an integral and systematic approach in a network-organisation (Rampersad, 2002).

In this paper knowledge is seen as a higher level of complexity than information, in the sense that it contains not only facts and insight in those facts, but also experiences, attitudes and skills. In this light is important to recognize that people know more than they can tell (Polanyi, 1962). Knowledge management includes the systematic planning, storing, controlling, using and distribution of knowledge that is important for the organisation as well as for the individuals in it (Weggeman, 2000). According to Feather (2004) one of the key aspects of knowledge management in an organisation is to ensure that the informal information that underpins effective operation is included. It is therefore essential that the explicit knowledge of an organisation is evaluated continually and managers play a crucial role in this process (Belasen, 2000).

Davenport & Prusak (1997) found that political battles in organisations could also frustrate the sharing of information. They claim that organisations need to take those aspects into consideration when a strategy for using and sharing information and knowledge is developed. This notion is supported by Rosenberg (2006) who also explains the importance of the culture of a learning organisation supporting the development and distribution of
knowledge in the organisation. Trust between staff members as well as staff members and management is important in a spirit of open communication, commitment and willingness to work together for a common goal. In this light the following quote of Drucker in Davenport & Prusak (1997:28) appears to be important: “We will have to learn, before understanding any task, to first ask the question, What information do I need, and in what form, and when?... The next question people have to learn to ask is, To whom do I owe which information and when and where.”

Communities of practice are communities that support professional discussion and work by sharing knowledge and experiences, often called best practices. Such communities often have some kind of online presence and sociability and usability are often important determining factors (Preece et al., 2004). Wenger (2000) argues that the success of organisations is largely determined by the ability to create communities of learning and practice.

**Competency management**

According to Nobre (2002) there are three types of competencies that are important in organisations: core competencies within an organisation, competencies that the employees have and competencies that the customers as the ultimate decision makers have.

According to Kluytmans (2005) and Harrison & Kessels (2004) it appears to be beneficial for an organisation to create harmony in which the potential human resources are optimally available for the organisation and at the same time the employability of the employee is increased. One of the adjusted roles of the manager is the need to find a balance between the competencies required by the organisation with the development needs of the staff he/she is responsible for (Kessels, 1999). The managers could identify the gaps in the required competencies and the available competencies and adjust development and recruitment programmes accordingly. With rapidly changing organisations employees often do not work within the strict limits of their function descriptions any longer (role-orientated), but the demands of the work increasingly requires them to be flexible and task-oriented (Kluytmans, 2005).

Electronic means are available to help in this regard (Stoof, 2005). A competency profile could be made and managed for each employee and core competencies in the organisation could be digitally stored and managed.

**ICT-security awareness management**

With improvements of ICT the global security environment has changed dramatically (English, 2005) and it has become necessary to deal with the risks in a fundamentally different and more effective way. It is not sufficient any longer to have security professionals in the organisation and technical solutions in place; Every employee needs to be aware of the ICT-security risks as well as the potential consequences of such security breaches (English, 2005; Siponen, 2001; Parker, 2001). Often not enough attention is given to this human aspect and as Siponen (2001:26) states: “Nothing is done as long as nothing goes wrong.” and yet the cost of doing nothing can be huge. Managers could play an essential role in making sure that every employee is aware of the ICT-security risks and thus improving the integrity, availability and exclusivity of information in the organisation. The manager needs to understand the importance of a high quality of information in terms of completeness and collective significance (English, 2005). Bonatti et al. (2006) argue that the most important cause of computer security violations on the Internet is the lack of technical knowledge of the users.

**Innovation and change management**

Based on research done by de Jong & den Hartog (2005) it could be argued that participation of employees in generating innovative ideas is essential and that strategic attention to innovation has a positive correlation with innovative behaviour of employees, thus emphasizing the role of the manager in this regard. Innovation is further enhanced if an organisation is able to attract talented and creative people of diverse backgrounds and able to create an open working climate (De Pree, 2006). A manager plays an important role in facilitating those aspects in the work environment.
In the dynamic environments of the information society it is important that managers coach their workers to see the importance of change so that they want to change themselves and want to make a positive contribution in the changes (Stoker, 2005). The importance of effective leadership and communication during continuing changes in organisations is recognized by a number of authors (Hargrove, 2001; Boonstra, 2005; Belasen, 2000; Stoker, 2005).

**Research method**

A conceptual framework for ICT and ICT-related competencies required by managers in the information society based on research by Broos (2007) was used to determine a topic list for the questionnaire as is illustrated in table 1. The questions in the questionnaire are based on this topic list. The number of questions for each aspect was mainly based on the number of topics identified for each theme. A pilot study was conducted by handing out questionnaires in various parts of the organisation to 15 members of the research population. The questionnaire was thereafter adjusted according to statistical findings and comments of the participants in the pilot.

<table>
<thead>
<tr>
<th>Table 1 Topic list for questionnaire based on research of Broos (2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topics identified</strong></td>
</tr>
<tr>
<td>General ICT skills</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Adjusted ways of working using ICT</td>
</tr>
<tr>
<td>Learning organisation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Knowledge management</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Communities of practice</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Attitude towards life long learning</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Competency management</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Focus on employees</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Change Management</td>
</tr>
<tr>
<td>Innovation management</td>
</tr>
</tbody>
</table>

A questionnaire was circulated amongst managers in a large non-profit organisation in the Netherlands (Broos, 2007). The first section of the questionnaire contained questions about the use of applications and general use of ICT in the work environment based on the topics mentioned in table 1. The use of ICT in the workplace contained a number of open questions about specific applications in the work environment. Furthermore a list of applications was stated in the questionnaire. The respondents could select one option on a 4-point scale for each application, where option one represents ‘do not use’, option two represents ‘not so important’, option three represents ‘important’ and option four represents ‘very important’. In order to obtain insight in whether the applications are used and their relative importance in the digital work environment the four options provided sufficient information.

The second section of the questionnaire contained statements about new ways of working in the information society as they were identified in the literature. The respondents could evaluate the statements on a 5-point Likert scale where option one represents ‘does not apply at all’ and option five represents ‘applies entirely’. The purpose of the second section of the questionnaire was to provide some insight in the commitment, knowledge and behaviour regarding the adjusted ways of working required in the information society by means of a self-evaluation and therefore a standard 5-point Likert scale with five ordered responses appeared appropriate for this section. The items included in the questionnaire for the new ways of working are listed below.

**Creating and participating in a learning organisation**

- I do communicate electronically with other professionals about my work.
- In my function it is important to continue to learn all the time.
- I have the opportunity to learn via the Internet during working hours.
- I obtain ideas from the work of others that I find on the Internet to improve my own work.
- I participate in keeping the information on the Intranet of the organisation up to date.
• I obtain ideas from the work of others that I find on the Intranet of the organisation to improve my own work.
• It is important to store the knowledge of my section electronically.
• I have the opportunity to learn via the Intranet of the organisation during working hours.
• I share my work-related knowledge with others electronically using a share.
• I spend time to organize electronically the working knowledge of the unit that I am responsible for.
• I think of ways to improve the sharing of information electronically.
• I spend time to improve the sharing of organisational knowledge electronically.
• I benefit from colleagues who share their experiences/lessons learnt with me.
• I share the mistakes that I made and what I learnt from it with my colleagues.
• I play an important role in managing the knowledge of the organisation electronically.
• I reflect on how information can be managed more effectively.
• I play an important role in organizing the flow of information in my unit.
• It is important for my organisation unit to share working knowledge and information with international partners.
• I implement new ways of working with information in the organisation.
• I reflect about the integrity of the information that I am responsible for.
• I allow my subordinates to learn via the Internet during working hours.
• I encourage my subordinates to share their working knowledge with others electronically.
• I discuss the advantages of sharing working knowledge electronically with my subordinates.

Competency management

• I use my computer to obtain insight in the competencies needed in the organisation.
• I use my computer to store relevant information about the potential of my subordinates.
• I recognize development needs of my subordinates.
• I facilitate the development needs of my subordinates.

ICT-security awareness

• I reflect on the security of information in the organisation.
• I know what the security risks of the Internet are.
• I reflect about the integrity of the information that I am responsible for.
• I encourage ICT-security awareness amongst my subordinates.

Change and innovation management

• I know how to manage change effectively.
• I know how I can accompany changes in the organisation effectively.
• Communication is important during a change in the organisation.
• I know how I can deal with the resistance my subordinates have against changes in the organisation.
• I consider renewal projects as a challenge.
• I have enough autonomy to work in the way I find best.
• Support in developing new ideas is always found in the organisation.
• I use creative ideas to improve the working method.
• I encourage my subordinates to participate in the thinking process about improving the working processes.
• I allow my subordinates to work in the way they find best.
• I allow my subordinates to make mistakes.

The questionnaire was sent to 700 randomly selected managers in the organisation. The managers have received a higher education, work at least one year in the organisation and are working in a variety of function areas ranging from governance, human resource management, administration and logistics, information and communication systems, planning and control, education and training to technical design and maintenance. 246 respondents returned
the questionnaire. This constitutes 35% of the sample. Relevant demographic variables of the respondents such as gender and position in the organisation compared sufficiently to the demographic variables of the target population in the organisation. The responses to the questionnaire were analyzed using the statistical package SPSS version 13. The respondents were given the opportunity to write comments on the questionnaire, which gave a further voice to the managers. Some of those comments are used to illustrate the experiences from an insider’s point of view. The comments were analyzed in a qualitative way by identifying predominant topics in the opinions of the respondents. Thereafter the comments related to each topic were compared to each other to obtain insight in the relative importance of the opinions about each topic. The identified topics were then compared and linked to the items identified for this research based on the literature review. The comments are used to enrich the discussion of the findings.

Finally factor analysis was used to further investigate if the separate components that are found in the literature review are also observed as separate components based on the analysis of the results of the questionnaire.

Findings

Managers spend on average 20.9 hours per week on the PC at work. Two of those hours are spent on the Internet and 3.8 of those hours on the Intranet of the organisation. Working at home during working hours is not yet common practice in this organisation with a mean of 1.1 hours per week (s.d. 3.8) but most managers indicated that the hours worked at home wereproductive.

In general the managers indicated that they were more productive as a result of the information and communication technology in their work environment, however 31% of the managers indicated that they lose production time because they are not familiar with the software applications in their work environment.

One respondent wrote on the questionnaire: “I observe a mixture of loose applications and an abundance of digitalization, but only poor automatization. The consequence is that it costs time instead of saves time”.

Managers are fairly confident in using ICT in their work environment, however

- 27% indicated that they are not able to use all the software applications that they needed in their work.
- 32% indicated problems managing e-mail effectively.
- 30% indicated problems managing information effectively. One respondent wrote in this regard on the questionnaire: “Information can easily be stored, but finding it afterwards is difficult”.
- 48% needed help from others in dealing with ICT in their work environment.

The following quote illustrates the frustration that some of the managers have regarding their lack of knowledge about ICT: “I have an ultra modern computer system with many recent applications, but I only use it as a typing and e-mail machine, since I have no insight in how to use the applications. I use pen, paper and the knowledge of my colleagues to do my work. I fall back on stone age technology since I am not familiar with the system.”

Managers did not experience regular info-stress, however

- 65% of the managers experience some stress as a result of using ICT (e.g. software, printers and availability of the network) in their work.
- 22% of the managers experience stress as a result of using ICT in their work because they do not have enough knowledge about it.
- 48% of the managers experience some stress as a result of the amount of information that they have to work through on a daily basis.

The next section discusses the importance that managers allocated to some of the applications in their work environment. Microsoft Word and Email were excluded from the list as these technologies can be regarded as ubiquitous. The rated applications are ordered according to perceived importance in table 2.

<table>
<thead>
<tr>
<th>Software application</th>
<th>Mean (Scale 1-4)</th>
<th>S.D.</th>
<th>95% Confidence Interval for Mean Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
</table>

Table 2 Response about importance of software applications in the working situation
Delving deeper into the data we can consider the extent to which managers find the applications important, use it, and ignore it.

- **Intranet**: 86.9% find it important, 11.4% use it but do not find it important and 1.6% do not use it.
- **MS Excel**: 76.8% find it important, 20.3% use it but do not find it important and 2.8% do not use it.
- **MS PowerPoint**: 76.4% find it important, 19.9% use it but do not find it important and 3.7% do not use it.
- **Electronic calendar**: 67% find it important, 11.4% use it but do not find it important and 20.8% do not use it.
- **Internet**: 60.4% find it important, 24.1% use it but do not find it important and 15.5% do not use it.
- **Information Management System**: 48.8% find it important, 17.4% use it but do not find it important and 33.3% do not use it.
- **MS Access**: 36% find it important, 35.3% use it but do not find it important and 28.7% do not use it.
- **Project Planning system**: 27.4% find it important, 20.3% use it but do not find it important and 52.3% do not use it.
- **Competency Management system**: 18.1% find it important, 15.2% use it but do not find it important and 66.7% do not use it.
- **Tools to organize thoughts**: 17% find it important, 13.7% use it but do not find it important and 69.3% do not use it.
- **Video conferencing**: 10.2% find it important, 4.1% use it but do not find it important and 84.5% do not use it.
- **On-line discussion**: 5.3% find it important, 10.2% use it but do not find it important and 84.5% do not use it.

The qualitative data showed that some managers simply did not know what certain applications could do for them: “There are applications that I use now and find very important. I would have liked to know about those applications at an earlier stage. An example is mind mapping. This is an important tool to organize thoughts. It is most likely important to all managers.”

In graph 1 is illustrated the number of years the respondents work as managers in the organisation. In this research no significant differences were found related to the number of years that managers were managing in the organisation in relation to the lack of knowledge.

This means that younger managers did not experience fewer problems using ICT in their work environment compared to older managers. This research result could be important since it is often expected that younger people will be able to deal with ICT in the work environment based on their experience in private use of the computer or what they have learnt in school. This research does not support such notion and therefore a re-evaluation of the curriculum of higher education in order to include more ICT related topics appears necessary. Especially in the light of higher education becoming increasingly needs-driven instead of content-driven (Plomp, 2006).

A number of managers indicated on the questionnaire that it was important to have a conceptual insight in the generic functionalities of the applications or in the context of exchanging data between applications. They wanted to understand the underlying principles of the applications, what the possibilities are of the applications and how applications can import and export data. Furthermore they wanted to know what the reasons are that there are sometimes limitations in doing so.
Some respondents wrote on the questionnaire that it is important to learn how to communicate appropriately using digital media, implying that this was not always the case in the organisation.

Table 3 shows how the aspects under consideration in this article applied to the respondents. The findings will be discussed in the following paragraph.

Table 3  The results of commitment, knowledge and behaviour to alternative ways of working for managers in the information society

<table>
<thead>
<tr>
<th>Component</th>
<th>Central tendency</th>
<th>95% confidence interval for mean</th>
<th>Cronbach’s alpha*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower bound</td>
<td>Upper bound</td>
</tr>
<tr>
<td>Creating and participating in a learning organisation</td>
<td>Applies partly</td>
<td>2.8</td>
<td>2.9</td>
</tr>
<tr>
<td>Competency management</td>
<td>Applies mainly</td>
<td>3.7</td>
<td>3.9</td>
</tr>
<tr>
<td>ICT-security awareness</td>
<td>Applies mainly</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td>Change and innovation management</td>
<td>Applies mainly</td>
<td>3.7</td>
<td>3.8</td>
</tr>
</tbody>
</table>

*Where scores of items are totalled to obtain a single score for a component, a Cronbach’s alpha was determined to ensure that the questions that contribute to those dimensions are acceptably homogeneous.

Discussion of commitment, knowledge and behaviour to alternative ways of working for managers

When the items related to the component ‘Participating in the learning organisation, knowledge management and communities of learning’ are combined in a scale, the managers score lower than expected, with a central tendency of ‘applies partly’. From the comments written on the questionnaires leadership emerged as an important factor, as this quotation typifies: “Leadership is the most important task of a manager; Working with people instead of hiding behind a computer.”

In connection with the learning organization three respondents wrote on the questionnaire that a cultural change was required in the organization in order to improve the sharing of knowledge and information, since in their opinion some managers might either not want to admit having made a mistake in order to avoid risking their careers or not want to share information and expertise from a viewpoint of power. A number of respondents indicated the importance of having a sound insight in the knowledge of the organisation, especially in the light of many re-organisations occurring simultaneously in their organisation.
One respondent indicated that managers needed to play a role preventing uncontrolled increase in information, stating "least is best". Respondents emphasized the importance of critically managing the knowledge of the organisation, evaluating the knowledge that is shared, keeping it up to date and removing outdated knowledge and information from a common database. A substantial amount of time is lost because managers cannot find relevant information when they need it. 38% of the managers waste time looking for information on the Internet and 67% of the managers waste time trying to find information on the Intranet of the organisation.

48% of the managers indicated that they experience difficulty evaluating the credibility of the information they find on the Internet.

Some managers indicated that using ICT in the work environment changes the culture of the organisation, which is shown by the following quote: "ICT is a good tool for the organisation, but at the same time a burden. ICT communication goes at the cost of interpersonal communication and personal interactions. The culture of the organisation is changing." One respondent stated that it is also important to have communities of practice in a wider context across the borders of a particular function area to share information and knowledge in order to improve integral management in the organisation.

Creating new rules in relation to ICT-security awareness is seen as one way to improve the security situation, but as one manager indicated: “To prohibit does not contribute to a culture of involvement. A better way is to work towards a culture of security in which the employees become 'security-minded'.” A number of respondents indicated that facilities needed to be created in order to make working possible in a secure environment without complicating the work.

General agreement exists amongst managers in this organisation about the continuing adjustments required in order to keep up with the changes regarding dealing effectively with information, communication and the technology.

Results of the factor analysis regarding alternative ways of working using ICT

Finally factor analysis is used to further investigate if the separate components that are found in the literature review can also be observed in the results of the questionnaire. Based on the results it could be determined if subcomponents need to be included and or certain components need to be grouped together. Factor analysis sorts the variables in homogeneous components based on the correlations they have with each other. Each component can thus explain a section of the variance of the results.

Six principal components are found with an initial Eigenvalue > 1 and a Cronbach’s alpha ≥7, explaining a total of 53% variance. Varimax-rotation is used to reorganize the information in a more effective way so that it is easier to interpret the components (ten Berge & Siero, 1997). The results of the factor analysis are summarized in table 4.

<table>
<thead>
<tr>
<th>Alternative ways of working</th>
<th>Explained variance %</th>
<th>Number of items</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating in the learning organisation and knowledge management in general</td>
<td>16</td>
<td>10</td>
<td>0.9</td>
</tr>
<tr>
<td>Innovation and change management</td>
<td>10</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>Knowledge management in own unit, including communication about it</td>
<td>8</td>
<td>6</td>
<td>0.8</td>
</tr>
<tr>
<td>Participating in communities of practice</td>
<td>7</td>
<td>3</td>
<td>0.7</td>
</tr>
<tr>
<td>ICT-security awareness</td>
<td>6</td>
<td>3</td>
<td>0.7</td>
</tr>
<tr>
<td>Competency management</td>
<td>6</td>
<td>3</td>
<td>0.7</td>
</tr>
</tbody>
</table>

From the table it appears as if participating in a learning organisation, knowledge management in own unit as well as participating in communities of practice could be seen as different components and hence it would be better to deal with them separately in further research. Innovation and change management, ICT-security awareness as well as competency management are identified as components.
Conclusions and recommendations

Although the results of this research can not be generalized it could provide a basis from which further research amongst managers in organisations could be conducted. Many of the results might be of a general nature and emphasize the importance of employers being aware of the influence of implementing and applying ICT in their organisations.

In conclusion we try to answer the two sub-questions, regarding existing and required skills, as stated in the introduction to this article.

What skills do they have already?

Managers use the company Intranet, MS Excel, MS PowerPoint, the electronic calendars and the Internet regularly in their work environment and find those applications important. Based on the comments written on the questionnaires it would be fair to suggest that, although they use these applications and find their use important, they often use those applications probably rather superficial. Nevertheless it can be safely assumed that expecting managers to communicate using MS PowerPoint and MS Excel, and expecting them to find information on the Internet and Intranet effectively, would be academically feasible.

More significantly, however, the feedback from the questionnaires indicates that the most important ICT-related competency that the managers seem to possess, is the desire to learn more, as well as the desire to participate in a learning organisation. It is our contention that this attitudinal quality actually outweighs the actual physical knowledge of how to manipulate a given interface or application. The most important item of knowledge that the managers already have, is the knowledge of their own shortcomings in terms of acquiring information from the Internet and from the corporate Intranet, shortcomings in processing that information adequately using existing software such as spreadsheets and databases; and finally they know they need to improve their ICT communication skills.

What ICT skills do the managers require?

From this research can be concluded that using software effectively in the context of the work situation is very important in order to function effectively. This is in accordance with the research results from den Boer & Hövels (2002). This research shows that although a number of managers appear to be ready to deal effectively with the information and communication technology as well as the adjusted ways of working and leading, a substantial number of managers experience problems because they do not have enough knowledge about software applications, hardware and networks and experience problems in using the applications effectively in their work environment.

What seems to emerge here however, is that the managers need to know more than just which buttons to push. They call for training that will lead to a conceptual understanding of ICT and its possibilities.

Managers spend a considerable amount of the work time using the PC. They play an important role in implementing ICT in the work environment and in motivating and coaching their staff in the adjusted ways of working. Our research indicates the importance of including effective use of ICT as well as dealing effectively with the adjusted ways of working in the corporate curriculum and perhaps also in higher education. In this regard it is remarkable that managers on average assign a relatively low score to the scale for ‘creating and participating in a learning organisation’ with the central tendency being ‘applies partly’. In the light of the important role of the manager in facilitating the alternative ways of working, this appears to indicate a need for further training or performance support. According to Rowlands (2003) information policy ought to be seen as a verb and not a noun, arguing that including such issues in a corporate curriculum is a continuous process and can never be a finished product. Harrison & Kessels (2004) argue that effective knowledge management will not happen automatically, but that HRD programmes should deal with those aspects. The results of this research support this notion.

It has also become clear that communication using technology is not per definition the same as communication through natural language (Hart-Davidson, 2001). A number of participants indicated that they are often not sure how to communicate using ICT. They indicated that they found it difficult to represent data effectively using the
technology. Once again there is a need for training, not so much in using ICT, but in understanding and working in an environment that is constantly affected by the ever-changing nature of ICT.

An important change encompasses the attitude to the ownership of information. This research found that some managers regard their experiences, knowledge and even information as their property and as such provide them with a sense of power. This provides some support for the ideas of Rosenberg (2006) and research findings of Davenport & Prusak (1997) about the importance of creating a culture conductive to sharing information and knowledge. This research shows some support for the notion of Brown & Duguid (2000) not to underestimate the need for social networks and resources as well as the need for personal contact in order to make learning and working possible.

The most important skill required by the managers, therefore, is the skill of surviving in an information-driven work environment. ICT literacy is a subset of the greater skill of acquiring, processing and communicating knowledge.

These findings could motivate a re-evaluation of curricula of higher educational institutes in order to prepare future managers and professionals to deal effectively with information, communication and ICT in their working practices.

References


